Falsification and the Methodology of Scientific Research Programmes[1]

IMRE LAKATOS
London School of Economics

1. *Science: reason or religion?*
2. *Fallibilism versus falsificationism.*
   (*a*) *Dogmatic (or naturalistic) falsificationism. The empirical basis.*
   (*b*) *Methodological falsificationism. The 'empirical basis'.*
   (*c*) *Sophisticated versus naive falsificationism. Progressive and degenerating problemshifts.*
3. *A methodology of scientific research programmes.*
   (*a*) *Negative heuristic; the 'hard core' of the programme.*
   (*b*) *Positive heuristic; the construction of the 'protective belt' and the relative autonomy of theoretical science.*
   (*c*) *Two illustrations: Prout and Bohr.*
      (*c1*) *Prout: a research programme progressing in an ocean of anomalies.*
      (*c2*) *Bohr: a research programme progressing on inconsistent foundations.*
   (*d*) *A new look at crucial experiments: the end of instant rationality.*
      (*d1*) *The Michelson–Morley experiment.*
      (*d2*) *The Lummer–Pringsheim experiments.*
      (*d3*) *Beta-decay versus conservation laws.*
      (*d4*) *Conclusion. The requirement of continuous growth.*
4. *The Popperian versus the Kuhnian research programme.*
Appendix: Popper, falsificationism and the 'Duhem–Quine thesis'.

1. SCIENCE: REASON OR RELIGION?

For centuries knowledge meant proven knowledge—proven either by the power of the intellect or by the evidence of the senses. Wisdom and intellectual integrity demanded that one must desist from unproven utterances and minimize, even in thought, the gap between speculation and established knowledge. The proving power of the intellect or the senses was

[1] This paper is a considerably improved version of my [1968*b*] and a crude version of my [1970]. Some parts of the former are here reproduced without change with the permission of the Editor of the *Proceedings of the Aristotelian Society*. In the preparation of the new version I received much help from Tad Beckman, Colin Howson, Clive Kilmister, Larry Laudan, Eliot Leader, Alan Musgrave, Michael Sukale, John Watkins and John Worrall.

questioned by the sceptics more than two thousand years ago; but they were browbeaten into confusion by the glory of Newtonian physics. Einstein's results again turned the tables and now very few philosophers or scientists still think that scientific knowledge is, or can be, proven knowledge. But few realize that with this the whole classical structure of intellectual values falls in ruins and has to be replaced: one cannot simply water down the ideal of proven truth—as some logical empiricists do—to the ideal of 'probable truth'[1] or—as some sociologists of knowledge do—to 'truth by [changing] consensus'.[2]

Popper's distinction lies primarily in his having grasped the full implications of the collapse of the best-corroborated scientific theory of all times: Newtonian mechanics and the Newtonian theory of gravitation. In his view virtue lies not in caution in avoiding errors, but in ruthlessness in eliminating them. Boldness in conjectures on the one hand and austerity in refutations on the other: this is Popper's recipe. Intellectual honesty does not consist in trying to entrench, or establish one's position by proving (or 'probabilifying') it—intellectual honesty consists rather in specifying precisely the conditions under which one is willing to give up one's position. Committed Marxists and Freudians refuse to specify such conditions: this is the hallmark of their intellectual dishonesty. *Belief* may be a regrettably unavoidable biological weakness to be kept under the control of criticism: but *commitment* is for Popper an outright crime.

Kuhn thinks otherwise. He too rejects the idea that science grows by accumulation of eternal truths.[3] He too takes his main inspiration from Einstein's overthrow of Newtonian physics. His main problem too is *scientific revolution*. But while according to Popper science is 'revolution in permanence', and criticism the heart of the scientific enterprise, according to Kuhn revolution is exceptional and, indeed, extra-scientific, and criticism is, in 'normal' times, anathema. Indeed for Kuhn the transition

[1] The main contemporary protagonist of the ideal of 'probable truth' is Rudolf Carnap. For the historical background and a criticism of this position, cf. Lakatos [1968*a*].

[2] The main contemporary protagonists of the ideal of 'truth by consensus' are Polanyi and Kuhn. For the historical background and a criticism of this position, cf. Musgrave [1969*a*], Musgrave [1969*b*] and Lakatos [1970].

[3] Indeed he introduces his [1962] by arguing against the 'development-by-accumulation' idea of scientific growth. But his intellectual debt is to Koyré rather than to Popper. Koyré showed that positivism gives bad guidance to the historian of science, for the history of physics can only be understood in the context of a succession of 'metaphysical' research programmes. Thus scientific changes are connected with vast cataclysmic metaphysical revolutions. Kuhn develops this message of Burtt and Koyré and the vast success of his book was partly due to his hard-hitting, direct criticism of justificationist historiography—which created a sensation among ordinary scientists and historians of science whom Burtt's, Koyré's (or Popper's) message has not yet reached. But, unfortunately, his message had some authoritarian and irrationalist overtones.

from criticism to commitment marks the point where progress—and 'normal' science—begins. For him the idea that on 'refutation' one can demand the rejection, the elimination of a theory, is 'naive' falsificationism. Criticism of the dominant theory and proposals of new theories are only allowed in the rare moments of 'crisis'. This last Kuhnian thesis has been widely criticized[1] and I shall not discuss it. My concern is rather that Kuhn, having recognized the failure both of justificationism and falsificationism in providing rational accounts of scientific growth, seems now to fall back on irrationalism.

For Popper scientific change is rational or at least rationally reconstructible and falls in the realm of the *logic of discovery*. For Kuhn scientific change—from one 'paradigm' to another—is a mystical conversion which is not and cannot be governed by rules of reason and which falls totally within the realm of the *(social) psychology of discovery*. Scientific change is a kind of religious change.

The clash between Popper and Kuhn is not about a mere technical point in epistemology. It concerns our central intellectual values, and has implications not only for theoretical physics but also for the underdeveloped social sciences and even for moral and political philosophy. If even in science there is no other way of judging a theory but by assessing the number, faith and vocal energy of its supporters, then this must be even more so in the social sciences: truth lies in power. Thus Kuhn's position would vindicate, no doubt, unintentionally, the basic political *credo* of contemporary religious maniacs ('student revolutionaries').

In this paper I shall first show that in Popper's logic of scientific discovery two different positions are conflated. Kuhn understands only one of these, 'naive falsificationism' (I prefer the term 'naive methodological falsificationism'); I think that his criticism of it is correct, and I shall even strengthen it. But Kuhn does not understand a more sophisticated position the rationality of which is not based on 'naive' falsificationism. I shall try to explain—and further strengthen—this stronger Popperian position which, I think, may escape Kuhn's strictures and present scientific revolutions as constituting rational progress rather than as religious conversions.

2. FALLIBILISM VERSUS FALSIFICATIONISM

(*a*) *Dogmatic (or naturalistic) falsificationism. The empirical basis.*

To see the conflicting theses more clearly, we have to reconstruct the problem situation as it was in philosophy of science after the breakdown of 'justificationism'.

[1] Cf. e.g. Watkins's and Feyerabend's contributions to this volume.

According to the 'justificationists' scientific knowledge consisted of proven propositions. Having recognized that strictly logical deductions enable us only to infer (transmit truth) but not to prove (establish truth), they disagreed about the nature of those propositions (axioms) whose truth can be proved by extra-logical means. *Classical intellectualists* (or 'rationalists' in the narrow sense of the term) admitted very varied—and powerful—sorts of extralogical 'proofs' by revelation, intellectual intuition, experience. These, with the help of logic, enabled them to prove every sort of scientific proposition. *Classical empiricists* accepted as axioms only a relatively small set of 'factual propositions' which expressed the 'hard facts'. Their truth-value was established by experience and they constituted the *empirical basis* of science. In order to prove scientific *theories* from nothing else but the narrow empirical basis, they needed a logic much more powerful than the deductive logic of the classical intellectualists: *'inductive logic'*. All justificationists, whether intellectualists or empiricists, agreed that a singular statement expressing a 'hard fact' may *disprove* a universal theory[1]; but few of them thought that a finite conjunction of factual propositions might be sufficient to *prove* 'inductively' a universal theory.[2]

Justificationism, that is, the identification of knowledge with proven knowledge, was the dominant tradition in rational thought throughout the ages. Scepticism did not deny justificationism: it only claimed that there was (and could be) no proven knowledge and *therefore* no knowledge whatsoever. For the sceptics 'knowledge' was nothing but animal belief. Thus justificationist scepticism ridiculed objective thought and opened the door to irrationalism, mysticism, superstition.

This situation explains the enormous effort invested by classical rationalists in trying to save the synthetical *a priori* principles of intellectualism and by classical empiricists in trying to save the certainty of an empirical basis and the validity of inductive inference. For all of them *scientific honesty demanded that one assert nothing that is unproven.* However, both were defeated: Kantians by non-Euclidean geometry and by non-Newtonian physics, and empiricists by the logical impossibility of establishing

[1] Justificationists repeatedly stressed this asymmetry between singular factual statements and universal theories. Cf. e.g. Popkin's discussion of Pascal in Popkin [1968], p. 14 and Kant's statement to the same effect as quoted in the new *motto* of the third 1969 German edition of Popper's *Logik der Forschung*. (Popper's choice of this time-honoured cornerstone of elementary logic as a *motto* of the new edition of his classic shows his main concern: to fight *probabilism*, in which this asymmetry becomes irrelevant; for probabilists theories may become almost as well established as factual propositions.)

[2] Indeed, even some of these few shifted, following Mill, the rather obviously insoluble problem of inductive proof (of universal from particular propositions) to the slightly less obviously insoluble problem of proving *particular* factual propositions from other *particular* factual propositions.

an empirical basis (as Kantians pointed out, facts cannot prove propositions) and of establishing an inductive logic (no logic can infallibly increase content). It turned out that *all theories are equally unprovable.*

Philosophers were slow to recognize this, for obvious reasons: classical justificationists feared that once they conceded that theoretical science is unprovable, they would have also to concede that it is sophistry and illusion, a dishonest fraud. The philosophical importance of *probabilism* (or *'neojustificationism'*) lies in the denial that such a concession is necessary.

Probabilism was elaborated by a group of Cambridge philosophers who thought that although scientific theories are equally unprovable, they have different degrees of probability (in the sense of the calculus of probability) relative to the available empirical evidence.[1] *Scientific honesty then requires less than had been thought: it consists in uttering only highly probable theories; or even in merely specifying, for each scientific theory, the evidence, and the probability of the theory in the light of this evidence.*

Of course, replacing proof by probability was a major retreat for justificationist thought. But even this retreat turned out to be insufficient. It was soon shown, mainly by Popper's persistent efforts, that under very general conditions all theories have zero probability, whatever the evidence; *all theories are not only equally unprovable but also equally improbable.*[2]

Many philosophers still argue that the failure to obtain at least a probabilistic solution of the problem of induction means that we 'throw over almost everything that is regarded as knowledge by science and common sense.'[3] It is against this background that one must appreciate the dramatic change brought about by falsificationism in evaluating theories, and in general, in the standards of intellectual honesty. Falsificationism was, in a sense, a new and considerable retreat for rational thought. But since it was a retreat from utopian standards, it cleared away much hypocrisy and muddled thought, and thus, in fact, it represented an advance.

First I shall discuss a most important brand of falsificationism: dogmatic (or 'naturalistic')[4] falsificationism. Dogmatic falsificationism admits the fallibility of *all* scientific theories without qualification, but it retains a sort of infallible empirical basis. It is strictly empiricist without being inductivist: it denies that the certainty of the empirical basis can be transmitted

[1] The founding fathers of probabilism were intellectualists; Carnap's later efforts to build up an empiricist brand of probabilism failed. Cf. my [1968*a*], p. 367 and also p. 361, footnote 2.

[2] For a detailed discussion, cf. my [1968*a*], especially pp. 353 ff.

[3] Russell [1943], p. 683. For a discussion of Russell's justificationism, cf. my [1962], especially pp. 167 ff.

[4] For the explanation of this term, cf. *below*, p. 98, footnote 1.

to theories. *Thus dogmatic falsificationism is the weakest brand of justificationism.*

It is extremely important to stress that admitting [fortified] empirical counterevidence as a final arbiter against a theory does not make one a dogmatic falsificationist. Any Kantian or inductivist will agree to such arbitration. But both the Kantian and the inductivist, while bowing to a negative crucial experiment, will also specify conditions of how to establish, entrench one unrefuted theory more than another. Kantians held that Euclidean geometry and Newtonian mechanics were established with certainty; inductivists held they had probability 1. For the dogmatic falsificationist, however, empirical *counter*evidence is the *one and only* arbiter which may judge a theory.

The hallmark of dogmatic falsificationism is then the recognition that all theories are equally conjectural. Science cannot *prove* any theory. But although science cannot *prove*, it can *disprove*: it 'can perform with complete logical certainty [the act of] repudiation of what is false',[1] that is, there is an absolutely firm empirical basis of facts which can be used to disprove theories. Falsificationists provide new—very modest—standards of scientific honesty: they are willing to regard a proposition as 'scientific' not only if it is a proven factual proposition, but even if it is nothing more than a falsifiable one, that is, if there are factual propositions available at the time with which it may clash, or, in other words, if it has potential falsifiers.[2]

Scientific honesty then consists of specifying, in advance, an experiment such that if the result contradicts the theory, the theory has to be given up.[3] The falsificationist demands that once a proposition is disproved, there must be no prevarication: the proposition must be unconditionally rejected. To (non-tautologous) unfalsifiable propositions the dogmatic falsificationist gives short shrift: he brands them 'metaphysical' and denies them scientific standing.

Dogmatic falsificationists draw a sharp demarcation between the theoretician and the experimenter: the theoretician proposes, the experimenter—in the name of Nature—disposes. As Weyl put it: 'I wish to record my unbounded admiration for the work of the experimenter in his struggle to wrest interpretable facts from an unyielding Nature who knows so well how to meet our theories with a decisive *No*—or with an inaudible

[1] Medawar [1967], p. 144.

[2] This discussion already indicates the vital importance of a demarcation between provable factual and unprovable theoretical propositions for the dogmatic falsificationist.

[3] '*Criteria of refutation* have to be laid down beforehand: it must be agreed which observable situations, if actually observed, mean that the theory is refuted' (Popper [1963], p. 38, footnote 3).

Yes.'[1] Braithwaite gives a particularly lucid exposition of dogmatic falsificationism. He raises the problem of the objectivity of science: 'To what extent, then, should an established scientific deductive system be regarded as a free creation of the human mind, and to what extent should it be regarded as giving an objective account of the facts of nature?'. His answer is: 'The form of a statement of a scientific hypothesis and its use to express a general proposition, is a human device; what is due to Nature are the observable facts which refute or fail to refute the scientific hypothesis . . . [In science] we hand over to Nature the task of deciding whether any of the contingent lowest-level conclusions are false. This objective test of falsity it is which makes the deductive system, in whose construction we have very great freedom, a deductive system of scientific hypotheses. Man proposes a system of hypotheses: Nature disposes of its truth or falsity. Man invents a scientific system, and then discovers whether or not it accords with observed fact.'[2]

According to the logic of dogmatic falsificationism, science grows by repeated overthrow of theories with the help of hard facts. For instance, according to this view, Descartes's vortex theory of gravity was refuted—and eliminated—by the *fact* that planets moved in ellipses rather than in Cartesian circles; Newton's theory, however, explained successfully the then available facts, both those which had been explained by Descartes's theory and those which refuted it. Therefore Newton's theory replaced Descartes's theory. Analogously, as seen by falsificationists, Newton's theory was, in turn, refuted—proved false—by the anomalous perihelion of Mercury, while Einstein's explained that too. Thus science proceeds by bold speculations, which are never proved or even made probable, but some of which are later eliminated by hard, conclusive refutations and then replaced by still bolder, new and, at least at the start, unrefuted speculations.

Dogmatic falsificationism, however, is untenable. It rests on two false assumptions and on a too narrow criterion of demarcation between scientific and non-scientific.

The *first assumption* is that there is a natural, *psychological* borderline between theoretical or speculative propositions on the one hand and factual or observational (or basic) propositions on the other. (I shall call this—following Popper—the *naturalistic doctrine of observation.*)

The *second assumption* is that if a proposition satisfies the psychological

[1] Quoted in Popper [1934], section 85, with Popper's comment: 'I fully agree.'

[2] Braithwaite [1953], pp. 367-8. For the 'incorrigibility' of Braithwaite's observed facts, cf. his [1938]. While in the quoted passage Braithwaite gives a forceful answer to the problem of scientific objectivity, in another passage he points out that 'except for the straightforward generalizations of observable facts . . . complete refutation is no more possible than is complete proof' ([1953], p. 19). Also cf. *below*, p. 113, footnote 4.

criterion of being factual or observational (or basic) then it is true; one may say that it was *proved* from facts. (I shall call this the *doctrine of observational (or experimental) proof.*)[1]

These two assumptions secure for the dogmatic falsificationist's deadly disproofs an empirical basis from which proven falsehood can be carried by deductive logic to the theory under test.

These assumptions are complemented by a *demarcation criterion*: only those theories are 'scientific' which forbid certain observable states of affairs and therefore are factually disprovable. *Or, a theory is 'scientific' if it has an empirical basis.*[2]

But both assumptions are false. Psychology testifies against the first, logic against the second, and, finally, methodological judgment testifies against the demarcation criterion. I shall discuss them in turn.

(1) A first glance at a few characteristic examples already undermines the *first assumption*. Galileo claimed that he could 'observe' mountains on the moon and spots on the sun and that these 'observations' refuted the time-honoured theory that celestial bodies are faultless crystal balls. But his 'observations' were not 'observational' in the sense of being observed by the—unaided—senses: their reliability depended on the reliability of his telescope—and of the optical theory of the telescope—which was violently questioned by his contemporaries. It was not Galileo's—pure, untheoretical—*observations* that confronted Aristotelian *theory* but rather Galileo's 'observations' in the light of his optical theory that confronted the Aristotelians' 'observations' in the light of their theory of the heavens.[3] This leaves us with two inconsistent theories, *prima facia* on a par. Some empiricists may concede this point and agree that Galileo's 'observations' were not genuine observations; but they still hold that there is a 'natural demarcation' between statements impressed on an empty and passive mind directly by the senses—only these constitute genuine 'immediate knowledge'—and between statements which are suggested by impure, theory-impregnated sensations. Indeed, *all* brands of justificationist theories of knowledge which acknowledge the senses as a source (whether as *one*

[1] For these assumptions and their criticism, cf. Popper [1934], sections 4 and 10. It is because of this assumption that—following Popper—I call this brand of falsificationism 'naturalistic'. Popper's 'basic propositions' should not be confused with the basic propositions discussed in this section; cf. *below*, p. 106, footnote 4.

It is important to point out that these two assumptions are also shared by many justificationists who are not falsificationists: they may add to experimental proofs 'intuitive proofs'—as did Kant—or 'inductive proofs'—as did Mill. Our falsificationist accepts experimental proofs *only*.

[2] The empirical basis of a theory is the set of its potential falsifiers: the set of those observational propositions which may disprove it.

[3] Incidentally, Galileo also showed—with the help of his optics—that if the moon was a faultless crystal ball, it would be invisible (Galileo [1632]).

source or as *the* source) of knowledge are bound to contain a *psychology of observation*. Such psychologies specify the 'right', 'normal', 'healthy', 'unbiased', 'careful' or 'scientific' state of the senses—or rather the state of mind as a whole—in which they observe truth as it is. For instance, Aristotle—and the Stoics—thought that the right mind was the medically healthy mind. Modern thinkers recognized that there is more to the right mind than simple 'health'. Descartes's right mind is one steeled in the fire of sceptical doubt which leaves nothing but the final loneliness of the *cogito* in which the *ego* can then be re-established and God's guiding hand found to recognize truth. All schools of modern justificationism can be characterized by the particular *psychotherapy* by which they propose to prepare the mind to receive the grace of proven truth in the course of a mystical communion. In particular, for classical empiricists the right mind is a *tabula rasa*, emptied of all original content, freed from all prejudice of theory. But it transpires from the work of Kant and Popper—and from the work of psychologists influenced by them—that such empiricist psychotherapy can never succeed. For there are and can be no sensations unimpregnated by expectations and therefore *there is no natural (i.e. psychological) demarcation between observational and theoretical propositions.*[1]

(2) But even if there was such a natural demarcation, logic would still destroy the *second assumption* of dogmatic falsificationism. For the truth-value of the 'observational' propositions cannot be indubitably decided: *no factual proposition can ever be proved from an experiment.* Propositions can only be derived from other propositions, they cannot be derived from facts: one cannot prove statements from experiences—'no more than by thumping the table.'[2] This is one of the basic points of elementary logic, but one which is understood by relatively few people even today.[3]

If factual propositions are unprovable then they are fallible. If they are fallible then clashes between theories and factual propositions are not 'falsifications' but merely inconsistencies. Our imagination may play a greater role in the formulation of 'theories' than in the formulation of

[1] True, most psychologists who turned against the idea of justificationist sensationalism did so under the influence of pragmatist philosophers like William James who denied the possibility of any sort of objective knowledge. But, even so, Kant's influence through Oswald Külpe, Franz Brentano and Popper's influence through Egon Brunswick and Donald Campbell played a role in the shaping of modern psychology; and if psychology ever vanquishes psychologism, it will be due to an increased understanding of the Kant-Popper mainline of objectivist philosophy.

[2] Cf. Popper [1934], section 29.

[3] It seems that the first philosopher to emphasize this might have been Fries in 1837 (cf. Popper [1934], section 29, footnote 3). This is of course a special case of the general thesis that logical relations, like probability or consistency, refer to *propositions*. Thus, for instance, the proposition 'nature is consistent' is false (or, if you wish, meaningless), for nature is not a proposition (or a conjunction of propositions).

'factual propositions',[1] but they are both fallible. Thus *we cannot prove theories and we cannot disprove them either*.[2] The demarcation between the soft, unproven 'theories' and the hard, proven 'empirical basis' is non-existent: *all* propositions of science are theoretical and, incurably, fallible.[3]

(3) Finally, even if there were a natural demarcation between observation statements and theories, and even if the truth-value of observation statements could be indubitably established, dogmatic falsificationism would still be useless for eliminating the most important class of what are commonly regarded as scientific theories. For even if experiments *could* prove experimental reports, their disproving power would still be miserably restricted: *exactly the most admired scientific theories simply fail to forbid any observable state of affairs.*

To support this last contention, I shall first tell a characteristic story and then propose a general argument.

The story is about an imaginary case of planetary misbehaviour. A physicist of the pre-Einsteinian era takes Newton's mechanics and his law of gravitation (N), the accepted initial conditions, I, and calculates, with their help, the path of a newly discovered small planet, p. But the planet deviates from the calculated path. Does our Newtonian physicist consider that the deviation was forbidden by Newton's theory and therefore that, once established, it refutes the theory N? No. He suggests that there must be a hitherto unknown planet p' which perturbs the path of p. He calculates the mass, orbit, etc., of this hypothetical planet and then asks an experimental astronomer to test his hypothesis. The planet p' is so small that even the biggest available telescopes cannot possibly observe it: the experimental astronomer applies for a research grant to build yet a bigger one.[4] In three years' time the new telescope is ready. Were the unknown planet p' to be discovered, it would be hailed as a new

[1] Incidentally, even this is questionable. Cf. *below*, pp. 127 ff.

[2] As Popper put it: 'No conclusive disproof of a theory can ever be produced'; those who wait for an infallible disproof before eliminating a theory will have to wait for ever and 'will never benefit from experience' ([1934], section 9).

[3] Kant and his English follower, Whewell, both realized that all scientific propositions, whether *a priori* or *a posteriori*, are equally theoretical; but both held that they are equally provable. Kantians saw clearly that the propositions of science are theoretical in the sense that they are not written by sensations on the *tabula rasa* of an empty mind, nor deduced or induced from such propositions. A factual proposition is only a special kind of theoretical proposition. In this Popper sided with Kant against the empiricist version of dogmatism. But Popper went a step further: in his view the propositions of science are not only theoretical but they are all also *fallible*, conjectural for ever.

[4] If the tiny conjectural planet were out of the reach even of the biggest *possible* optical telescopes, he might try some quite novel instrument (like a radiotelescope) in order to enable him to 'observe it', that is, to ask Nature about it, even if only indirectly. (The new 'observational' theory may itself not be properly articulated, let alone severely tested, but he would care no more than Galileo did.)

victory of Newtonian science. But it is not. Does our scientist abandon Newton's theory and his idea of the perturbing planet? No. He suggests that a cloud of cosmic dust hides the planet from us. He calculates the location and properties of this cloud and asks for a research grant to send up a satellite to test his calculations. Were the satellite's instruments (possibly new ones, based on a little-tested theory) to record the existence of the conjectural cloud, the result would be hailed as an outstanding victory for Newtonian science. But the cloud is not found. Does our scientist abandon Newton's theory, together with the idea of the perturbing planet and the idea of the cloud which hides it? No. He suggests that there is some magnetic field in that region of the universe which disturbed the instruments of the satellite. A new satellite is sent up. Were the magnetic field to be found, Newtonians would celebrate a sensational victory. But it is not. Is this regarded as a refutation of Newtonian science? No. Either yet another ingenious auxiliary hypothesis is proposed or . . . the whole story is buried in the dusty volumes of periodicals and the story never mentioned again.[1]

This story strongly suggests that even a most respected scientific theory, like Newton's dynamics and theory of gravitation, may fail to forbid any observable state of affairs.[2] Indeed, *some scientific theories forbid an event occurring in some specified finite spatio-temporal region (or briefly, a 'singular event') only on the condition that no other factor* (possibly hidden in some distant and unspecified spatio-temporal corner of the universe) *has any influence on it.* But then *such theories never alone contradict a 'basic' statement:* they contradict at most a conjunction of a basic statement describing a spatio-temporally singular event and of a universal non-existence statement saying that no other relevant cause is at work anywhere in the universe. And the dogmatic falsificationist cannot possibly claim that such universal non-existence statements belong to the empirical basis: that they can be observed and proved by experience.

Another way of putting this is to say that some scientific theories are normally interpreted as containing a *ceteris paribus* clause[3]: in such cases it is always a specific theory *together* with this clause which may be refuted. But such a refutation is inconsequential for the *specific* theory under test

[1] At least not until a new research programme supersedes Newton's programme which happens to explain this previously recalcitrant phenomenon. In this case, the phenomenon will be unearthed and enthroned as a 'crucial experiment'; cf. *below*, pp. 154 ff.

[2] Popper asks: 'What kind of clinical responses would refute to the satisfaction of the analyst not merely a particular diagnosis but psychoanalysis itself?' ([1963], p. 38, footnote 3.) But what kind of observation would refute to the satisfaction of the Newtonian not merely a particular version but Newtonian theory itself?

[3] [*Added in press:*] This '*ceteris paribus*' clause must not normally be interpreted as a separate premise. For a discussion, cf. *below*, p. 186.

because by replacing the *ceteris paribus* clause by a different one the *specific* theory can always be retained whatever the tests say.

If so, the 'inexorable' disproof procedure of dogmatic falsificationism breaks down in these cases *even if* there were a firmly established empirical basis to serve as a launching pad for the arrow of the *modus tollens*: the prime target remains hopelessly elusive.[1] And as it happens, it is exactly the most important, 'mature' theories in the history of science which are *prima facie* undisprovable in this way.[2] Moreover, by the standards of dogmatic falsificationism all probabilistic theories also come under this head: for no finite sample can ever *disprove* a universal probabilistic theory[3]; probabilistic theories, like theories with a *ceteris paribus* clause, have no empirical basis. But then the dogmatic falsificationist relegates the most important scientific theories *on his own admission* to metaphysics where rational discussion—consisting, by his standards, of proofs and disproofs—has no place, since a metaphysical theory is neither provable nor disprovable. The demarcation criterion of dogmatic falsificationism is thus still strongly antitheoretical.

(Moreover, *one can easily argue that ceteris paribus clauses are not exceptions, but the rule in science.* Science, after all, must be demarcated from a curiosity shop where funny local—or cosmic—oddities are collected and displayed. The assertion that 'all Britons died from lung cancer between 1950 and 1960' is logically possible, and might even have been true. But if it has been only an occurrence of an event with minute probability, it would have only curiosity value for the crankish fact-collector, it would have a macabre entertainment value, but no scientific value. A proposition might be said to be scientific only if it aims at expressing a causal connection: such connection between being a Briton and dying of lung cancer may not even be intended. Similarly, 'all swans are white', if true, would be a mere curiosity unless it asserted that swanness *causes* whiteness. But then a black swan would not refute this proposition, since it may only indicate *other causes* operating simultaneously. Thus 'all swans are white' is either an oddity and easily disprovable or a scientific proposition with a *ceteris paribus* clause and therefore undisprovable. *Tenacity of a theory against empirical evidence would then be an argument for rather than against regarding it as 'scientific'. 'Irrefutability' would become a hallmark of science.*[4])

[1] Incidentally, we might persuade the dogmatic falsificationist that his demarcation criterion was a very naive mistake. If he gives it up but retains his two basic assumptions, he will have to ban theories from science and regard the growth of science as an accumulation of proven basic statements. This indeed is the final stage of classical empiricism after the evaporation of the hope that facts can prove or at least disprove theories.

[2] This is no coincidence; cf. *below*, pp. 175 ff.

[3] Cf. Popper [1934], chapter VIII.

[4] For a *much* stronger case, cf. *below*, sect. 3.

To sum up: classical justificationists only admitted proven theories; neoclassical justificationists probable ones; dogmatic falsificationists realized that in either case no theories are admissible. They decided to admit theories if they are disprovable—disprovable by a finite number of observations. But even if there were such disprovable theories—those which can be contradicted by a finite number of observable facts—they are still logically too near to the empirical basis. For instance, on the terms of the dogmatic falsificationist, a theory like 'All planets move in ellipses' may be disproved by five observations; therefore the dogmatic falsificationist will regard it as scientific. A theory like 'All planets move in circles' may be disproved by four observations; therefore the dogmatic falsificationist will regard it as still more scientific. The acme of scientificness will be a theory like 'All swans are white' which is disprovable by one single observation. On the other hand, he will reject all probabilistic theories together with Newton's, Maxwell's, Einstein's theories, as unscientific, for no finite number of observations can ever disprove them.

If we accept the demarcation criterion of dogmatic falsificationism, *and* also the idea that facts can prove 'factual' propositions, we have to declare that the most important, if not all, theories ever proposed in the history of science are metaphysical, that most, if not all, of the accepted progress is pseudo-progress, that most, if not all, of the work done is irrational. If, however, still accepting the demarcation criterion of dogmatic falsificationism, we deny that facts can prove propositions, then we certainly end up in complete scepticism: then all science is undoubtedly irrational metaphysics and should be rejected. *Scientific theories are not only equally unprovable, and equally improbable, but they are also equally undisprovable.* But the recognition that not only the theoretical but *all* the propositions in science are fallible, means the total collapse of *all* forms of dogmatic justificationism as theories of scientific rationality.

(*b*) *Methodological falsificationism. The 'empirical basis'.*

The collapse of dogmatic falsificationism because of fallibilistic arguments seems to bring us back to square one. If *all* scientific statements are fallible theories, one can criticize them only for inconsistency. But then, in what sense, if any, is science empirical? If scientific theories are neither provable, nor probabilifiable, nor disprovable, then the sceptics seem to be finally right: science is no more than vain speculation and there is no such thing as progress in scientific knowledge. Can we still oppose scepticism? *Can we save scientific criticism from fallibilism?* Is it possible to have a fallibilistic theory of scientific progress? In particular, if scientific criticism is fallible, on what ground can we ever eliminate a theory?

A most intriguing answer is provided by *methodological falsificationism*. Methodological falsificationism is a brand of conventionalism; therefore in order to understand it, we must first discuss conventionalism in general.

There is an important demarcation between *'passivist' and 'activist' theories of knowledge*. 'Passivists' hold that true knowledge is Nature's imprint on a perfectly inert mind: mental *activity* can only result in bias and distortion. The most influential passivist school is classical empiricism. 'Activists' hold that we cannot read the book of Nature without mental activity, without interpreting them in the light of our expectations or theories.[1] Now *conservative 'activists'* hold that we are born with our basic expectations; with them we turn the world into 'our world' but must then live for ever in the prison of our world. The idea that we live and die in the prison of our 'conceptual frameworks' was developed primarily by Kant; pessimistic Kantians thought that the real world is for ever unknowable because of this prison, while optimistic Kantians thought that God created our conceptual framework to fit the world.[2] But *revolutionary activists* believe that conceptual frameworks can be developed and also replaced by new, *better* ones; it is *we* who create our 'prisons' and we can also, critically, demolish them.[3]

New steps from conservative to revolutionary activism were made by Whewell and then by Poincaré, Milhaud and Le Roy. Whewell held that theories are developed by trial and error—in the 'preludes to the inductive epochs'. The best ones among them are then 'proved'—during the 'inductive epochs'—by a long primarily *a priori* consideration which he called 'progressive intuition'. The 'inductive epochs' are followed by 'sequels to the inductive epochs': cumulative developments of auxiliary theories.[4] Poincaré, Milhaud and Le Roy were averse to the idea of *proof* by progressive intuition and preferred to explain the continuing historical success of Newtonian mechanics by a *methodological decision* taken by scientists: after a considerable period of initial empirical success scientists

[1] This demarcation—and terminology—is due to Popper; cf. especially his [1934], section 19 and his [1945], chapter 23 and footnote 3 to chapter 25.

[2] No version of conservative activism explained why Newton's *gravitational* theory should be invulnerable; Kantians restricted themselves to the explanation of the tenacity of Euclidean geometry and Newtonian *mechanics*. About Newtonian *gravitation* and *optics* (or other branches of science) they had an ambiguous, and occasionally inductivist position.

[3] I do not include Hegel among 'revolutionary *activists*'. For Hegel and his followers change in conceptual frameworks is a predetermined, inevitable process, where individual creativity or rational criticism plays no essential role. Those who run ahead are equally at fault as those who stay behind in this 'dialectic'. The clever man is not he who creates a better 'prison' or who demolishes critically the old one, but the one who is always in step with history. Thus dialectic accounts for change without criticism.

[4] Cf. Whewell's [1837], [1840] and [1858].

may *decide* not to allow the theory to be refuted. Once they have taken this decision, they solve (or dissolve) the apparent anomalies by auxiliary hypotheses or other 'conventionalist stratagems.'[1] This *conservative conventionalism* has, however, the disadvantage of making us unable to get out of our self-imposed prisons, once the first period of trial-and-error is over and the great decision taken. It cannot solve the problem of the elimination of those theories which have been triumphant for a long period. According to conservative conventionalism, experiments may have sufficient power to refute young theories, but not to refute old, established theories: *as science grows, the power of empirical evidence diminishes.*[2]

Poincaré's critics refused to accept his idea, that, although the scientists build their conceptual frameworks, there comes a time when these frameworks turn into prisons which cannot be demolished. This criticism gave rise to two rival schools of *revolutionary conventionalism*: Duhem's simplicism and Popper's methodological falsificationism.[3]

Duhem accepts the conventionalists' position that no physical theory ever crumbles merely under the weight of 'refutations', but claims that it still may crumble under the weight of 'continual repairs, and many tangled-up stays' when 'the worm-eaten columns' cannot support 'the tottering building' any longer[4]; then the theory loses its original simplicity and has to be replaced. But falsification is then left to subjective taste or, at best, to scientific fashion, and leaves too much leeway for dogmatic adherence to a favourite theory.[5]

Popper set out to find a criterion which is both more objective and more

[1] Cf. especially Poincaré [1891] and [1902]; Milhaud [1896]; Le Roy [1899] and [1901]. It was one of the chief philosophical merits of conventionalists to direct the limelight to the fact that any theory can be saved by 'conventionalist stratagems' from refutations. (The term 'conventionalist stratagem' is Popper's; cf. the critical discussion of Poincaré's conventionalism in his [1934], especially sections 19 and 20.)

[2] Poincaré first elaborated his conventionalism only with regard to geometry (cf. his [1891]). Then Milhaud and Le Roy generalized Poincaré's idea to cover all branches of accepted physical theory. Poincaré's [1902] starts with a strong criticism of the Bergsonian Le Roy against whom he defends the empirical (falsifiable or 'inductive') character of all physics *except for* geometry and mechanics. Duhem, in turn, criticized Poincaré: in his view there was a possibility of overthrowing even Newtonian mechanics.

[3] The *loci classici* are Duhem's [1905] and Popper's [1934]. Duhem was not a *consistent* revolutionary conventionalist. Very much like Whewell, he thought that conceptual changes are only *preliminaries* to the final—if perhaps distant—'natural classification': 'The more a theory is perfected, the more we apprehend that the logical order in which it arranges experimental laws is the reflection of an ontological order.' In particular, he refused to see Newton's mechanics *actually* 'crumbling' and characterized Einstein's relativity theory as the manifestation of a 'frantic and hectic race in pursuit of a novel idea' which 'has turned physics into a real chaos where logic loses its way and common-sense runs away frightened' (Preface—of 1914—to the second edition of his [1905]).

[4] Duhem [1905], chapter VI, section 10.

[5] For a further discussion of conventionalism, cf. *below*, pp. 184-189.

hard-hitting. He could not accept the emasculation of empiricism, inherent even in Duhem's approach, and proposed a methodology which allows experiments to be powerful even in 'mature' science. Popper's methodological falsificationism is both conventionalist and falsificationist, but he 'differs from the [conservative] conventionalists in holding that the statements decided by agreement are *not* [spatio-temporally] universal but [spatio-temporally] singular'[1]; and he differs from the dogmatic falsificationist in holding that the truth-value of such statements cannot be proved by facts but, in some cases, may be decided by agreement.[2]

The *conservative conventionalist* (or methodological justificationist, if you wish) makes unfalsifiable by *fiat* some (spatio-temporally) universal theories, which are distinguished by their explanatory power, simplicity or beauty. Our *revolutionary conventionalist* (or 'methodological falsificationist') makes unfalsifiable by *fiat* some (spatio-temporally) singular statements which are distinguishable by the fact that there exists at the time a 'relevant technique' such that 'anyone who has learned it' will be able to *decide* that the statement is 'acceptable'.[3] Such a statement may be called an 'observational' or 'basic' statement, but only in inverted commas.[4] Indeed, the very selection of all such statements is a matter of a decision, which is not based on exclusively psychological considerations. This decision is then followed by a second kind of decision concerning the separation of the set of *accepted* basic statements from the rest.

These *two decisions* correspond to the *two assumptions* of dogmatic falsificationism. But there are important differences. First, the methodological falsificationist is not a justificationist, he has no illusions about 'experimental proofs' and is fully aware of the fallibility of his decisions and the risks he is taking.

The methodological falsificationist realizes that in the 'experimental techniques' of the scientist fallible theories are involved,[5] 'in the light of which' he interprets the facts. In spite of this he 'applies' these theories, he regards them in the given context not as theories under test but as *unproblematic background knowledge* 'which we accept (tentatively) as unproblematic while we are testing the theory'.[6] He may call these theories —and the statements whose truth-value he decides in their light—'observational': but this is only a manner of speech which he inherited from naturalistic falsificationism.[1] The methodological falsificationist *uses our most successful theories as extensions of our senses* and widens the range of theories which can be applied in testing far beyond the dogmatic falsificationist's range of strictly observational theories. For instance, let us imagine that a big radio-star is discovered with a system of radio-star satellites orbiting it. We should like to test some gravitational theory on this planetary system—a matter of considerable interest. Now let us imagine that Jodrell Bank succeeds in providing a set of space-time co-ordinates of the planets which is inconsistent with the theory. We shall take these statements as potential falsifiers. Of course, these basic statements are not 'observational' in the usual sense but only " 'observational' ". They describe planets that neither the human eye nor optical instruments can reach. Their truth-value is arrived at by an 'experimental technique'. This 'experimental technique' is based on the 'application' of a well-corroborated theory of radio-optics. Calling these statements 'observational' is no more than a manner of saying that, in the context of his problem, that is, in testing our gravitational theory, the methodological falsificationist uses radio-optics uncritically, as 'background knowledge'. *The need for decisions to demarcate the theory under test from unproblematic background knowledge is a characteristic feature of this brand of methodological falsificationism.*[2] (This situation does not really differ from Galileo's 'observation' of Jupiter's satellites: moreover, as some of Galileo's contemporaries rightly pointed out, he relied on a virtually non-existent optical theory—which then was less corroborated, and even less articulated, than present-day radio-optics. On the other hand, calling the reports of our human eye 'observational' only indicates that we 'rely' on some vague physiological theory of human vision.[3])

This consideration shows the conventional element in granting—in a given context—the (methodologically) 'observational' status to a theory.[4] Similarly, there is a considerable conventional element in the decision concerning the actual truth-value of a basic statement which we take after we have decided which 'observational theory' to apply. One single observation may be the stray result of some trivial error: in order to reduce such risks, methodological falsificationists prescribe some safety control. The simplest such control is to repeat the experiment (it is a matter of

Popper [1934], section 30.

[2] *In this section I discuss the 'naive' variant of Popper's methodological falsificationism. Thus, throughout the section 'methodological falsificationism' stands for 'naive methodological falsificationism'; for this 'naivety', cf. below,* pp. 115-116.

[3] Popper [1934], section 27.

[4] *Op. cit.* section 28. For the non-basicness of these methodologically 'basic' statements, cf. e.g. Popper [1934] *passim* and Popper [1959*a*], p. 35, footnote *2.

[5] Cf. Popper [1934], end of section 26 and also his [1968*c*], pp. 291–2.

[6] Cf. Popper [1963], p. 390.

[1] Indeed, Popper carefully puts 'observational' in quotes; cf. his [1934], section 28.

[2] This demarcation plays a role both in the *first* and in the *fourth* type of decisions of the methodological falsificationist. (For the *fourth* decision, cf. *below*, p. 110.)

[3] For a fascinating discussion, cf. Feyerabend [1969].

[4] One wonders whether it would not be better to make a break with the terminology of naturalistic falsificationism and rechristen observational theories '*touchstone theories*'.

convention how many times); another is to 'fortify' the potential falsifier by a 'well-corroborated falsifying hypothesis'.[1]

The methodological falsificationist also points out that, as a matter of fact, these conventions are institutionalized and endorsed by the scientific community; the list of 'accepted' falsifiers is provided by the verdict of the experimental scientists.[2]

This is how the methodological falsificationist establishes his 'empirical basis'. (He uses inverted commas in order 'to give ironical emphasis' to the term.[3]) This 'basis' can be hardly called a 'basis' by justificationist standards: there is nothing proven about it—it denotes 'piles driven into a swamp'.[4] Indeed, if this 'empirical basis' clashes with a theory, the theory may be *called* 'falsified', but it is not falsified in the sense that it is disproved. Methodological 'falsification' is very different from dogmatic falsification. If a theory is falsified, it is proven false; if it is 'falsified', it may still be true. If we follow up this sort of 'falsification' by the actual 'elimination' of a theory, we may well end up by eliminating a true, and accepting a false, theory (a possibility which is thoroughly abhorrent to the old-fashioned justificationist).

Yet the methodological falsificationist advises that exactly this is to be done. The methodological falsificationist realizes that if we want to reconcile fallibilism with (non-justificationist) rationality, we *must* find a way to eliminate *some* theories. If we do not succeed, the growth of science will be nothing but growing chaos.

Therefore the methodological falsificationist maintains that '[if we want] to make the method of selection by elimination work, and to ensure that only the fittest theories survive, their struggle for life must be made severe'.[5] Once a theory has been falsified, in spite of the risk involved, it must be eliminated: '[with theories we work only] as long as they stand up to tests'.[6] The elimination must be methodologically conclusive: 'In general we regard an inter-subjectively testable falsification as final . . . A corroborative appraisal made at a later date . . . can replace a positive degree of corroboration by a negative one, but not *vice versa*'.[7] This is the

[1] Cf. Popper [1934], section 22. Many philosophers overlooked Popper's important qualification that a basic-statement has no power to refute anything without the support of a well-corroborated falsifying hypothesis.

[2] Cf. Popper [1934], section 30.

[3] Popper [1963], p. 387.

[4] Popper [1934], section 30; also cf. section 29: 'The Relativity of Basic Statements.'

[5] Popper [1957], p. 134. Popper, in other places, emphasizes that his method cannot 'ensure' the survival of the fittest. Natural selection may go wrong: the fittest may perish and monsters survive.

[6] Popper [1935].

[7] Popper [1934], section 82.

methodological falsificationist's explanation of how we get out of a rut: 'It is always the experiment which saves us from following a track that leads nowhere.'[1]

The methodological falsificationist separates rejection and disproof, which the dogmatic falsificationist had conflated.[2] He is a fallibilist but his fallibilism does not weaken his critical stance: he turns fallible propositions into a 'basis' for a hard-line policy. On these grounds he proposes a *new demarcation criterion*: only those theories—that is, non-'observational' propositions—which forbid certain 'observable' states of affairs, and therefore may be 'falsified' and rejected, are 'scientific': or, briefly, *a theory is 'scientific' (or 'acceptable') if it has an 'empirical basis'*. This criterion brings out sharply the difference between dogmatic and methodological falsificationism.[3]

This methodological demarcation criterion is much more liberal than the dogmatic one. Methodological falsificationism opens up new avenues of criticism: many more theories may qualify as 'scientific'. We have already seen that there are more 'observational' theories than observational theories,[4] and therefore there are more 'basic' statements than basic statements.[5] Furthermore, probabilistic theories may qualify now as 'scientific': although they are not falsifiable they can be easily made 'falsifiable' by an *additional (third type) decision* which the scientist can make by specifying certain rejection rules which may make statistically interpreted evidence 'inconsistent' with the probabilistic theory.[6]

[1] Popper [1934], section 82.

[2] This kind of methodological 'falsification' is, unlike dogmatic falsification (disproof), a pragmatic, methodological idea. But then what exactly should we mean by it? Popper's answer—which I am going to discard—is that methodological 'falsification' indicates an 'urgent need of replacing a falsified hypothesis by a better one' (Popper [1959*a*], p. 87, footnote *1). This shift is an excellent illustration of the process I described in my [1963–4] whereby critical discussion shifts the original *problem* without necessarily changing the old *terms*. The byproducts of such processes are *meaning-shifts*. For a further discussion, cf. *below*, p. 122, footnote 4, and p. 157, footnote 1.

[3] The demarcation criterion of the dogmatic falsificationist was: a theory is 'scientific' if it has an empirical basis (see *above*, p. 98).

[4] See *above*, pp. 98–9.

[5] Incidentally, Popper, in his [1934], does not seem to have seen this point clearly. He writes: 'Admittedly, it is possible to interpret the concept of an *observable event* in a psychologistic sense. But I am using it in such a sense that it might just as well be replaced by "an event involving position and movement of macroscopic physical bodies" '. ([1934], section 28.) In the light of our discussion, for instance, we may regard a positron passing through a Wilson chamber at time t_0 as an 'observable' event, in spite of the non-macroscopic character of the positron.

[6] Popper [1934], section 68. Indeed, this methodological falsificationism is the philosophical basis of some of the most interesting developments in modern statistics. The Neyman-Pearson approach rests completely on methodological falsificationism. Also cf. Braithwaite [1953], chapter VI. (Unfortunately, Braithwaite reinterprets Popper's demarcation criterion as separating meaningful from meaningless rather than scientific from non-scientific propositions.)

But even these three decisions are not sufficient to enable us to 'falsify' a theory which cannot explain anything 'observable' without a *ceteris paribus* clause.[1] No finite number of 'observations' is enough to 'falsify' such a theory. However, if this is the case how can one reasonably defend a methodology which claims to 'interpret natural laws or theories as . . . statements which are partially decidable, i.e. which are, for logical reasons, not verifiable but, in an asymmetrical way, falsifiable . . .'?[2] How can we interpret theories like Newton's theory of dynamics and gravitation as 'one-sidedly decidable'?[3] How can we make in such cases genuine 'attempts to weed out false theories—to find the weak points of a theory in order to reject it if it is falsified by the test'?[4] How can we draw them into the realm of rational discussion? The methodological falsificationist solves the problem by making a further (*fourth type*) *decision*: when he tests a theory together with a *ceteris paribus* clause and finds that this conjunction has been refuted, he must decide whether to take the refutation also as a refutation of the specific theory. For instance, he may accept Mercury's 'anomalous' perihelion as a refutation of the treble conjunction N_3 of Newton's theory, the known initial conditions and the *ceteris paribus* clause. Then he tests the initial conditions 'severely'[5] and may decide to relegate them into the 'unproblematic background knowledge'. This decision implies the refutation of the double conjunction N_2 of Newton's theory and the *ceteris paribus* clause. Now he has to take the crucial decision: whether to relegate also the *ceteris paribus* clause into the pool of 'unproblematic background knowledge'. He will do so if he finds the *ceteris paribus* clause well corroborated.

How can one test a *ceteris paribus* clause severely? By assuming that there *are* other influencing factors, by specifying such factors, and by testing these specific assumptions. If many of them are refuted, the *ceteris paribus* clause will be regarded as well-corroborated.

Yet the decision to 'accept' a *ceteris paribus* clause is a very risky one because of the grave consequences it implies. If it is decided to accept it as part of such background knowledge, the statements describing Mercury's perihelion from the empirical basis of N_2 are turned into the empirical basis of Newton's specific theory N_1 and what was previously a mere 'anomaly' in relation to N_1, becomes now crucial evidence against it, its falsification. (We may call an event described by a statement A an '*anomaly* in relation to a theory T' if A is a potential falsifier of the conjunction of T and a *ceteris paribus* clause but it becomes a potential falsifier of T itself after

[1] Cf. *above*, pp. 101–3.
[2] Popper [1933]. [3] Popper [1933]. [4] Popper [1957], p. 133.
[5] For a discussion of this important concept of Popperian methodology, cf. my [1968*a*], pp. 397 ff.

having decided to relegate the *ceteris paribus* clause into 'unproblematic background knowledge.'[1]) Since, for our savage falsificationist, falsifications are methodologically conclusive,[2] the fateful decision amounts to the methodological elimination of Newton's theory, making further work on it irrational. If the scientist shrinks back from such bold decisions he will 'never benefit from experience', 'believing, perhaps, that it is his business to defend a successful system against criticism as long as it is not *conclusively disproved*'.[3] He will degenerate into an apologist who may always claim that 'the discrepancies which are asserted to exist between the experimental results and the theory are only apparent and that they will disappear with the advance of our understanding'.[4] But for the falsificationist this is 'the very reverse of the critical attitude which is the proper one for the scientist',[5] and is impermissible. To use one of the methodological falsificationist's favourite expressions: the theory 'must be made to stick its neck out'.

The methodological falsificationist is in a serious plight when it comes to deciding where to draw the demarcation, even if only in a well-defined context, between the problematic and unproblematic. The plight is most dramatic when he has to make a decision about *ceteris paribus* clauses, when he has to promote one of the hundreds of 'anomalous phenomena' into a 'crucial experiment', and decide that in such a case the experiment was 'controlled'.[6]

Thus, with the help of this fourth type of decision,[7] our methodological falsificationist has finally succeeded in interpreting even theories like Newton's theory as 'scientific'.[8]

[1] For an improved 'explication', cf. *below*, p. 159, footnote 1.
[2] Cf. *above*, p. 108, text to footnotes 6 and 7.
[3] Popper [1934], section 9.
[4] *Ibid.* [5] *Ibid.*
[6] The problem of '*controlled experiment*' may be said to be nothing else but the problem of arranging experimental conditions in such a way as to minimize the risk involved in such decisions.
[7] This type of decision belongs, in an important sense, to the same category as the first decision: it demarcates, by decision, problematic from unproblematic knowledge. Cf. *above*, p. 107, text to footnote 2.
[8] Our exposition shows clearly the complexity of the decisions needed to define the 'empirical content' of a theory—that is, the set of its potential falsifiers. 'Empirical content' depends on our *decision* as to which are our 'observational theories' and which anomalies are to be promoted to counterexamples. If one attempts to compare the empirical content of different scientific theories in order to see which is 'more scientific', then one will get involved in an enormously complex and therefore hopelessly arbitrary system of decisions about their respective classes of 'relatively atomic statements' and their 'fields of application'. (For the meaning of these (very) technical terms, cf. Popper [1934], section 38.) But such comparison is possible only when one theory supersedes another (cf. Popper, [1959*a*], p. 401, footnote 7). And even then, there may be difficulties (which would not, however, add up to irremediable 'incommensurability').

Indeed, there is no reason why he should not go yet another step. Why not decide that a theory—which even these four decisions cannot turn into an empirically falsifiable one—is falsified if it clashes with another theory which is scientific on some of the previously specified grounds and is also well-corroborated?[1] After all, if we reject one theory because one of its potential falsifiers is seen to be true in the light of an observational theory, why not reject another theory because it clashes *directly* with one that may be relegated into unproblematic background knowledge? This would allow us, by a *fifth type decision*, to eliminate even 'syntactically metaphysical' theories, that is, theories, which, like 'all-some' statements or purely existential statements,[2] because of their *logical form* cannot have spatio-temporally singular potential falsifiers.

To sum up: the methodological falsificationist offers an interesting solution to the problem of combining hard-hitting criticism with fallibilism. Not only does he offer a philosophical basis for falsification after fallibilism had pulled the carpet from under the feet of the dogmatic falsificationist, but he also widens the range of such criticism very considerably. By putting falsification in a new setting, he saves the attractive code of honour of the dogmatic falsificationist: that scientific honesty consists in specifying, in advance, an experiment such, that if the result contradicts the theory, the theory has to be given up.[3]

Methodological falsificationism represents a considerable advance beyond both dogmatic falsificationism and conservative conventionalism. It recommends risky decisions. But the risks are daring to the point of recklessness and one wonders whether there is no way of lessening them.

Let us first have a closer look at the risks involved.

Decisions play a crucial role in this methodology—as in any brand of conventionalism. Decisions however may lead us disastrously astray. The methodological falsificationist is the first to admit this. But this, he argues, is the price which we have to pay for the possibility of progress.

One has to appreciate the dare-devil attitude of our methodological falsificationist. He feels himself to be a hero who, faced with two catastrophic alternatives, dared to reflect coolly on their relative merits and choose the lesser evil. One of the alternatives was sceptical fallibilism, with its 'anything goes' attitude, the despairing abandonment of all intellectual standards, and hence of the idea of scientific progress. Nothing can be established, nothing can be rejected, nothing even communicated: the

[1] This was suggested by J. O. Wisdom: cf. his [1963].

[2] For instance: 'All metals have a solvent'; or 'There exists a substance which can turn all metals into gold'. For discussions of such theories, cf. especially Watkins [1957] and Watkins [1960]. But cf. *below*, pp. 126–7 and pp. 183–4.

[3] See *above*, p. 96.

growth of science is a growth of chaos, a veritable Babel. For two thousand years, scientists and scientifically-minded philosophers chose justificationist illusions of some kind to escape this nightmare. Some of them argued that *one has to choose between inductivist justificationism and irrationalism*: 'I do not see any way out of a dogmatic assertion that we know the inductive principle or some equivalent; the only alternative is to throw over almost everything that is regarded as knowledge by science and common sense.'[1] Our methodological falsificationist proudly rejects such escapism: he dares to measure up to the full impact of fallibilism and yet escape scepticism by a daring and risky conventionalist policy, with no dogmas. He is fully aware of the risks but insists that *one has to choose between some sort of methodological falsificationism and irrationalism*. He offers a game in which one has little hope of winning, but claims that it is still better to play than to give up.[2]

Indeed, those critics of naive falsificationism who offer no alternative method of criticism are inevitably driven to irrationalism. For instance, Neurath's muddled argument, that the falsification and ensuing elimination of a hypothesis may turn out to have been 'an obstacle in the progress of science',[3] carries no weight as long as the only alternative he seems to offer is chaos. Hempel is, no doubt, right in stressing that 'science offers various examples [when] a conflict between a highly-confirmed theory and an occasional recalcitrant experiential sentence may well be resolved by revoking the latter rather than by sacrificing the former'[4]; nevertheless he admits that he can offer no other 'fundamental standard' than that of naive falsificationism.[5] Neurath—and, seemingly, Hempel—reject falsificationism as 'pseudo-rationalism'[6]; but where is 'real rationalism'? Popper warned already in 1934 that Neurath's permissive methodology (or rather lack of methodology) would make science unempirical and therefore

[1] Russell [1943], p. 683.

[2] I am sure that some will welcome methodological falsificationism as an 'existentialist' philosophy of science.

[3] Neurath [1935], p. 356.

[4] Hempel [1952], p. 621. Agassi, in his [1966], follows Neurath and Hempel, especially pp. 16 ff. It is rather amusing that Agassi, in making this point, thinks that he is taking up arms against 'the whole literature concerning the methods of science'.

Indeed, many scientists were fully aware of the difficulties inherent in the 'confrontation of theory and facts'. (Cf. Einstein [1949], p. 27.) Several philosophers sympathetic to falsificationism emphasized that 'the process of refuting a scientific hypothesis is more complicated than it appears to be at first sight' (Braithwaite [1953], p. 20). But only Popper offered a constructive, rational solution.

[5] Hempel [1952], p. 622. Hempel's crisp 'theses on empirical certainty' do nothing but refurbish Neurath's—and some of Popper's—old arguments (against Carnap, I take it); but deplorably, he does not mention either his predecessors or his adversaries.

[6] Neurath [1935].

irrational: 'We need a set of rules to limit the arbitrariness of "deleting" (or else "accepting") a protocol sentence. Neurath fails to give any such rules and thus unwittingly throws empiricism overboard . . . Every system becomes defensible if one is allowed (as everybody is, in Neurath's view) simply to "delete" a protocol sentence if it is inconvenient'.[1] Popper agrees with Neurath that all propositions are fallible; but he forcefully makes the crucial point that we cannot make progress unless we have a firm rational strategy or method to guide us when they clash.[2]

But is not the firm strategy of the brand of methodological falsificationism hitherto discussed *too firm*? Are not the decisions it advocates bound to be *too arbitrary*? Some may even claim that all that distinguishes methodological from dogmatic falsificationism is that *it pays lip-service to fallibilism*!

To criticize a theory of criticism is usually very difficult. Naturalistic falsificationism was relatively easy to refute, since it rested on an empirical psychology of perception: one could show that it was simply *false*. But how can methodological falsificationism be falsified? No disaster can ever disprove a non-justificationist theory of rationality. Moreover, how can we ever recognize an epistemological disaster? We have no means to judge whether the verisimilitude of our successive theories increases or decreases.[3] At this stage we have not yet developed a general theory of criticism even for scientific theories, let alone for theories of rationality[4]: therefore if we want to falsify our methodological falsificationism, we have to do it before having a theory of how to do it.

If we look at history of science, if we try to see how some of the most celebrated falsifications happened, we have to come to the conclusion that either some of them are plainly irrational, or that they rest on rationality principles radically different from the ones we just discussed. First of all, our falsificationist must deplore the fact that stubborn theoreticians frequently challenge experimental verdicts and have them reversed. In the falsificationist conception of scientific 'law and order' we have described there is no place for such successful appeals. Further difficulties arise from the falsification of theories to which a *ceteris paribus* clause is appended.[5]

[1] Popper [1934], section 26.

[2] Neurath's [1935] shows that he never grasped Popper's simple argument.

[3] I am using here 'verisimilitude' in Popper's sense: the difference between the truth content and falsity content of a theory. For the risks involved in estimating it, cf. my [1968*a*], especially pp. 395 ff.

[4] I tried to develop such a general theory of criticism in my [1970].

[5] The falsification of theories depends on the high degree of corroboration of the *ceteris paribus* clause. This however is not always the case. This is why the methodological falsificationist may advise us to rely on our 'scientific instinct' (Popper [1934], section 18, footnote 2) or 'hunch' (Braithwaite [1953], p. 20).

Their falsification as it occurs in actual history is *prima facie* irrational by the standards of our falsificationist. By his standards, scientists frequently seem to be irrationally slow: for instance, eighty-five years elapsed between the acceptance of the perihelion of Mercury as an anomaly and its acceptance as a falsification of Newton's theory, in spite of the fact that the *ceteris paribus* clause was reasonably well corroborated. On the other hand, scientists frequently seem to be irrationally rash: for instance, Galileo and his disciples accepted Copernican heliocentric celestial mechanics in spite of the abundant evidence against the rotation of the Earth; or Bohr and his disciples accepted a theory of light emission in spite of the fact that it ran counter to Maxwell's well-corroborated theory.

Indeed, it is not difficult to see at least two crucial characteristics common to both dogmatic and our methodological falsificationism which are clearly dissonant with the actual history of science: that (1) *a test is—or must be made—a two-cornered fight between theory and experiment so that in the final confrontation only these two face each other; and* (2) *the only interesting outcome of such confrontation is* (*conclusive*) *falsification*: '[*the only genuine*] *discoveries are refutations of scientific hypotheses.*'[1] However, history of science suggests that (1′) tests are—at least—three-cornered fights between rival theories and experiment and (2′) some of the most interesting experiments result, *prima facie*, in confirmation rather than falsification.

But if—as seems to be the case—the history of science does not bear out our theory of scientific rationality, we have two alternatives. One alternative is to abandon efforts to give a rational explanation of the success of science. Scientific method (or 'logic of discovery'), conceived as the discipline of rational appraisal of scientific theories—and of criteria of *progress*—vanishes. We, may, of course, still try to explain *changes* in 'paradigms' in terms of social psychology.[2] This is Polanyi's and Kuhn's way.[3] The other alternative is to try at least to *reduce* the conventional element in falsificationism (we cannot possibly eliminate it) and replace the

[1] Agassi [1959]; he calls Popper's idea of science '*scientia negativa*' (Agassi [1968]).

[2] It should be mentioned here that the Kuhnian sceptic is still left with what I would call the '*scientific sceptic's dilemma*': any scientific sceptic will still try to explain changes in beliefs and will regard his own psychology as a theory which is more than simple belief, which, in some sense, is 'scientific'. Hume, while trying to show up science as a mere system of beliefs with the help of his stimulus-response theory of learning, never raised the problem of whether his theory of learning applies also to his own theory of learning. In contemporary terms, we might well ask, does the popularity of Kuhn's philosophy indicate that people recognize its *truth*? In this case it would be refuted. Or does this popularity indicate that people regarded it as an attractive new fashion? In this case, it would be 'verified'. But would Kuhn like *this* 'verification'?

[3] Feyerabend who contributed probably more than anybody else to the spread of Popper's ideas, seems now to have joined the enemy camp. Cf. his intriguing [1970].

naive versions of methodological falsificationism—characterized by the theses (1) and (2) above—by a *sophisticated* version which would give a new *rationale* of falsification and thereby rescue methodology and the idea of scientific *progress*. This is Popper's way, and the one I intend to follow.

(*c*) *Sophisticated versus naive methodological falsificationism. Progressive and degenerating problemshifts.*

Sophisticated falsificationism differs from naive falsificationism both in its rules of *acceptance* (or 'demarcation criterion') and its rules of *falsification* or elimination. For the naive falsificationist any theory which can be interpreted as experimentally falsifiable, is 'acceptable' or 'scientific'.[1] For the sophisticated falsificationist a theory is 'acceptable' or 'scientific' only if it has corroborated excess empirical content over its predecessor (or rival), that is, only if it leads to the discovery of novel facts. This condition can be analysed into two clauses: that the new theory has excess empirical content ('*acceptability*'$_1$) and that some of this excess content is verified ('*acceptability*'$_2$). The first clause can be checked instantly[2] by *a priori* logical analysis; the second can be checked only empirically and this may take an indefinite time.

Again, for the naive falsificationist a theory is *falsified* by a '(fortified'[3]) 'observational' statement which conflicts with it (or rather, which he decides to interpret as conflicting with it). The sophisticated falsificationist regards a scientific theory T as falsified if and only if another theory T' has been proposed with the following characteristics: (1) T' has excess empirical content over T: that is, it predicts *novel* facts, that is, facts improbable in the light of, or even forbidden, by T;[4] (2) T' explains the previous success of T, that is, all the unrefuted content of T is contained (within the limits of observational error) in the content of T'; and (3) some of the excess content of T' is corroborated.[5]

In order to be able to appraise these definitions we need to understand their problem background and their consequences. First, we have to remember the conventionalists' methodological discovery that no experimental result can ever kill a theory: any theory can be saved from counterinstances either by some auxiliary hypothesis or by a suitable reinterpretation of its terms. Naive falsificationists solved this problem by relegating—in crucial contexts—the auxiliary hypotheses to the realm of unproblematic background knowledge, eliminating them from the deductive model of the test-

[1] Cf. *above*, p. 109. [2] But *cf. below*, pp. 155–7. [3] Cf. *above*, p. 108, text to footnote 1.

[4] I use 'prediction' in a wide sense that includes 'postdiction'.

[5] *For a detailed discussion of these acceptance and rejection rules and for references to Popper's work*, cf. my [1968*a*], pp. 375–90. For some qualifications (concerning continuity and consistency as regulative principles), cf. *below*, pp. 131–2 and 141–6.

situation and thereby *forcing* the chosen theory into logical isolation, in which it becomes a sitting target for the attack of test-experiments. But since this procedure did not offer a suitable guide for a rational reconstruction of the history of science, we may just as well completely rethink our approach. Why aim at falsification at any price? Why not rather impose certain standards on the theoretical adjustments by which one is allowed to save a theory? Indeed, some such standards have been well-known for centuries, and we find them expressed in age-old wisecracks against *ad hoc* explanations, empty prevarications, face-saving, linguistic tricks.[1] We have already seen that Duhem adumbrated such standards in terms of 'simplicity' and 'good sense'.[2] But *when* does lack of 'simplicity' in the protective belt of theoretical adjustments reach the point at which the theory *must* be abandoned?[3] In what sense was Copernican theory, for instance, 'simpler' than Ptolemaic?[4] The vague notion of Duhemian 'simplicity' leaves, as the naive falsificationist correctly argued, the decision very much to taste and fashion.[5]

Can one improve on Duhem's approach? Popper did. His solution—a sophisticated version of methodological falsificationism—is more objective and more rigorous. Popper agrees with the conventionalists that theories and factual propositions can always be harmonized with the help of auxiliary hypotheses: he agrees that the problem is how to demarcate between scientific and pseudoscientific *adjustments*, between rational and irrational changes of theory. According to Popper, saving a theory with the help of auxiliary hypotheses which satisfy certain well-defined conditions represents scientific progress; but saving a theory with the help of auxiliary hypotheses which do not, represents degeneration. Popper calls such inadmissible auxiliary hypotheses *ad hoc* hypotheses, mere linguistic devices, 'conventionalist stratagems'.[6] But then any scientific theory has to

[1] Molière, for instance, ridiculed the doctors of his *Malade Imaginaire*, who offered the *virtus dormitiva* of opium as the answer to the question as to why opium produced sleep. One might even argue that Newton's famous dictum *hypotheses non fingo* was really directed against *ad hoc* explanations—like his own explanation of gravitational forces by an aether-model in order to meet Cartesian objections. [2] Cf. *above*, p. 105.

[3] Incidentally, Duhem agreed with Bernard that experiments alone—without simplicity considerations—can decide the fate of theories in physiology. But in physics, he argued, they cannot ([1905], chapter VI, section 1).

[4] Koestler correctly points out that only Galileo created the myth that the Copernican theory was simple (Koestler [1959], p. 476); in fact, 'the motion of the earth [had not] done much to simplify the old theories, for though the objectionable equants had disappeared, the system was still bristling with auxiliary circles' (Dreyer [1906], chapter XIII). [5] Cf. *above*, p. 105.

[6] Popper [1934], sections 19 and 20. I have discussed in some detail—under the heads 'monster-barring', 'exception-barring', 'monster-adjustment'—such stratagems as they appear in informal, quasi-empirical mathematics; cf. my [1963–4].

be appraised together with its auxiliary hypotheses, initial conditions, etc., and, especially, together with its predecessors so that we may see by what sort of *change* it was brought about. Then, of course, what we appraise is a *series of theories* rather than isolated *theories*.

Now we can easily understand why we formulated the criteria of acceptance and rejection of sophisticated methodological falsificationism as we did.[1] But it may be worth while to reformulate them slightly, couching them explicitly in terms of *series of theories*.

Let us take a series of theories, T_1, T_2, T_3, ... where each subsequent theory results from adding auxiliary clauses to (or from semantical reinterpretations of) the previous theory in order to accommodate some anomaly, each theory having at least as much content as the unrefuted content of its predecessor. Let us say that such a series of theories is *theoretically progressive* (*or 'constitutes a theoretically progressive problemshift'*) if each new theory has some excess empirical content over its predecessor, that is, if it predicts some novel, hitherto unexpected fact. Let us say that a theoretically progressive series of theories is also *empirically progressive* (*or 'constitutes an empirically progressive problemshift'*) if some of this excess empirical content is also corroborated, that is, if each new theory leads us to the actual discovery of some *new fact*.[2] Finally, let us call a problemshift *progressive* if it is both theoretically and empirically progressive, and *degenerating* if it is not.[3] We '*accept*' problemshifts as 'scientific' only if they are at least theoretically progressive; if they are not, we '*reject*' them as 'pseudoscientific'. Progress is measured by the degree to which a problemshift is progressive, by the degree to which the series of theories leads us to the discovery of novel facts. We regard a theory in the series 'falsified' when it is superseded by a theory with higher corroborated content.[4]

This demarcation between progressive and degenerating problemshifts sheds new light on the appraisal of *scientific—or, rather, progressive—*

[1] Cf. *above*, p. 116.

[2] If I already know P_1: 'Swan A is white', P_ω: 'All swans are white' represents no progress, because it may only lead to the discovery of such further similar facts as P_2: 'Swan B is white'. So-called 'empirical generalizations' constitute no progress. A *new* fact must be improbable or even impossible in the light of previous knowledge. Cf. *above*, p. 116, and *below*, pp. 155 ff.

[3] The appropriateness of the term 'problemshift' for a series of theories rather than of problems may be questioned. I chose it partly because I have not found a more appropriate alternative—'theoryshift' sounds dreadful—partly because theories are always problematical, they never solve all the problems they have set out to solve. Anyway, in the second half of the paper, the more natural term 'research programme' will replace 'problemshifts' in the most relevant contexts.

[4] For the 'falsification' of certain series of theories (of 'research programmes') as opposed to the 'falsification' of one theory within the series, cf. *below*, pp. 155 ff.

explanations. If we put forward a theory to resolve a contradiction between a previous theory and a counterexample in such a way that the new theory, instead of offering a content-increasing (scientific) ***explanation***, only offers a content-decreasing (linguistic) ***reinterpretation***, the contradiction is resolved in a merely semantical, unscientific way. ***A given fact is explained scientifically only if a new fact is also explained with it.***[1]

Sophisticated falsificationism thus shifts the problem of how to appraise *theories* to the problem of how to appraise *series of theories*. Not an isolated *theory*, but only a series of theories can be said to be scientific or unscientific: to apply the term 'scientific' to one *single* theory is a category mistake.[2]

The time-honoured empirical criterion for a satisfactory theory was agreement with the observed facts. Our empirical criterion for a series of theories is that it should produce new facts. ***The idea of growth and the concept of empirical character are soldered into one.***

This revised form of methodological falsificationism has many new features. First, it denies that 'in the case of a scientific theory, our decision depends upon the results of experiments. If these confirm the theory, we may accept it until we find a better one. If they contradict the theory, we reject it.'[3] It denies that 'what ultimately decides the fate of a theory is the result of a test, *i.e.* an agreement about basic statements'.[4] Contrary to naive falsificationism, ***no experiment, experimental report, observation statement or well-corroborated low-level falsifying hypothesis alone can lead to falsification.***[5] ***There is no falsification before the emergence of a better theory.***[6]

[1] Indeed, in the original manuscript of my [1968*a*] I wrote: 'A theory without excess corroboration has no excess explanatory power; ***therefore, according to Popper, it does not represent growth and therefore it is not "scientific"; therefore, we should say, it has no explanatory power***' (p. 386). I cut out the italicized half of the sentence under pressure from my colleagues who thought it sounded too eccentric. I regret it now.

[2] Popper's conflation of 'theories' and 'series of theories' prevented him from getting the basic ideas of sophisticated falsificationism across more successfully. His ambiguous usage led to such confusing formulations as 'Marxism [as the core of a series of theories or of a "research programme"] is irrefutable' and, at the same time, 'Marxism [as a particular conjunction of this core and some specified auxiliary hypotheses, initial conditions and a *ceteris paribus* clause] has been refuted.' (Cf. Popper [1963].)

Of course, there is nothing wrong in saying that an isolated, single theory is 'scientific' if it represents an advance on its predecessor, as long as one clearly realizes that in this formulation we appraise the theory as the outcome of—and in the context of—a certain historical development.

[3] Popper [1945], vol. II, p. 233. Popper's more sophisticated attitude surfaces in the remark that 'concrete and practical consequences can be *more* directly tested by experiment' (*ibid.* my italics).

[4] Popper [1934], section 30.

[5] For the *pragmatic* character of methodological 'falsification', cf. *above*, p. 109, footnote 2.

[6] 'In most cases we have, before falsifying a hypothesis, another one up our sleeves' (Popper [1959*a*], p. 87, footnote *1). But, as our argument shows, we *must* have one. Or, as Feyerabend put it: 'The best criticism is provided by those theories which can replace the rivals they have removed' ([1965], p. 227). He notes that in *some* cases 'alternatives

But then the distinctively negative character of naive falsificationism vanishes; criticism becomes more difficult, and also positive, constructive. But, of course, if falsification depends on the emergence of better theories, on the invention of theories which anticipate new facts, then falsification is *not* simply a relation between a theory and the empirical basis, but a multiple relation between competing theories, the original 'empirical basis', and the empirical growth resulting from the competition. Falsification can thus be said to have a '*historical character*'.[1] Moreover, some of the theories which bring about falsification are frequently proposed *after* the 'counterevidence'. This may sound paradoxical for people indoctrinated with naive falsificationism. Indeed, this epistemological theory of the relation between theory and experiment differs sharply from the epistemological theory of naive falsificationism. The very term 'counterevidence' has to be abandoned in the sense that no experimental result must be interpreted directly as 'counterevidence'. If we still want to retain this time-honoured term, we have to redefine it like this: 'counterevidence to T_1' is a corroborating instance to T_2 which is either inconsistent with or independent of T_1 (with the *proviso* that T_2 is a theory which satisfactorily explains the empirical success of T_1). This shows that '*crucial counterevidence*'—or '*crucial experiments*'—can be recognized as such among the scores of anomalies only *with hindsight*, in the light of some superseding theory.[2]

Thus the crucial element in falsification is whether the *new theory* offers any novel, excess information compared with its predecessor and whether some of this excess information is corroborated. Justificationists valued 'confirming' instances of a theory; naive falsificationists stressed 'refuting' instances; for the methodological falsificationists it is the—rather rare—corroborating instances of the *excess* information which are the crucial ones; these receive all the attention. We are no longer interested in the

---

will be quite indispensable for the purpose of refutation' (*ibid.* p. 254). But according to our argument *refutation without an alternative shows nothing but the poverty of our imagination in providing a rescue hypothesis.* Also cf. *below*, p. 121, footnote 4.

[1] Cf. my [1968*a*], pp. 387 ff.

[2] In the distorting mirror of naive falsificationism, new theories which replace old refuted ones, are themselves born unrefuted. Therefore they do not believe that there is a relevant difference between anomalies and crucial counterevidence. For them, anomaly is a dishonest euphemism for counterevidence. But in actual history new theories are born refuted: they inherit many anomalies of the old theory. Moreover, frequently it is *only* the new theory which dramatically predicts that fact which will function as crucial counterevidence against its predecessor, while the 'old' anomalies may well stay on as 'new' anomalies.

All this will be still clearer when we introduce the idea of 'research programme': cf. *below*, pp. 135 and 176 ff.

thousands of trivial verifying instances nor in the hundreds of readily available anomalies: the few crucial *excess-verifying instances* are decisive.[1] This consideration rehabilitates—and reinterprets—the old proverb: *Exemplum docet, exempla obscurant.*

'Falsification' in the sense of naive falsificationism (corroborated counterevidence) is not a *sufficient* condition for eliminating a specific theory: in spite of hundreds of known anomalies we do not regard it as falsified (that is, eliminated) until we have a better one.[2] Nor is 'falsification' in the naive sense *necessary* for falsification in the sophisticated sense: a progressive problemshift does not have to be interspersed with 'refutations'. Science can grow without any 'refutations' leading the way. Naive falsificationists suggest a linear growth of science, in the sense that theories are followed by powerful refutations which eliminate them; these refutations in turn are followed by new theories.[3] It is perfectly *possible* that theories be put forward 'progressively' in such a rapid succession that the 'refutation' of the n-th appears only as the corroboration of the $n+1$-th. The problem fever of science is raised by proliferation of rival theories rather than counterexamples or anomalies.

This shows that the slogan *of proliferation of theories* is much more important for sophisticated than for naive falsificationism. For the naive falsificationist science grows through repeated experimental overthrow of theories; new rival theories proposed before such 'overthrows' may speed up growth but are not absolutely necessary[4]; constant proliferation of theories is optional but not mandatory. For the sophisticated falsificationist

[1] *Sophisticated falsificationism adumbrates a new theory of learning;* cf. *below*, p. 123.

[2] It is clear that the theory T' may have excess corroborated empirical content over another theory T even if both T and T' are refuted. Empirical content has nothing to do with truth or falsity. Corroborated contents can also be compared irrespective of the refuted content. Thus we may see the rationality of the elimination of Newton's theory in favour of Einstein's, even though Einstein's theory may be said to have been born—like Newton's —'refuted'. We have only to remember that 'qualitative confirmation' is a euphemism for 'quantitative disconfirmation'. (Cf. my [1968*a*], pp. 384–6.)

[3] Cf. Popper [1934], section 85, p. 279 of the 1959 English translation.

[4] It is true that a certain type of *proliferation of rival theories* is allowed to play an accidental heuristic role in falsification. In many cases falsification heuristically 'depends on [the condition] that sufficiently many and sufficiently different theories are offered' (Popper [1940]). For instance, we may have a theory T which is apparently unrefuted. But it may happen that a new theory T', inconsistent with T, is proposed which equally fits the available facts: the differences are smaller than the range of observational error. In such cases the inconsistency prods us into improving our 'experimental techniques', and thus refining the 'empirical basis' so that either T or T' (or, incidentally, both) can be falsified: 'We need [a] new theory in order to find out where the old theory was deficient' (Popper [1963], p. 246). But the role of this proliferation is *accidental* in the sense that, once the empirical basis is refined, the fight is between this refined empirical basis and the theory T under test; the rival theory T' acted only as a *catalyst*. (Also cf. *above*, p. 119, footnote 6.)

proliferation of theories cannot wait until the accepted theories are 'refuted' (or until their protagonists get into a Kuhnian crisis of confidence).[1] While naive falsificationism stresses 'the urgency of replacing a *falsified* hypothesis by a better one',[2] sophisticated falsificationism stresses the urgency of replacing *any* hypothesis by a better one. Falsification cannot 'compel the theorist to search for a better theory',[3] simply because falsification cannot precede the better theory.

The problem-shift from naive to sophisticated falsificationism involves a semantic difficulty. For the naive falsificationist a 'refutation' is an experimental result which, by force of his decisions, is made to conflict with the theory under test. But according to sophisticated falsificationism one must not take such decisions before the alleged 'refuting instance' has become the confirming instance of a new, better theory. Therefore whenever we see terms like 'refutation', 'falsification', 'counterexample', we have to check in each case whether these terms are being applied in virtue of decisions by the naive or by the sophisticated falsificationist.[4]

Sophisticated methodological falsificationism offers new standards for intellectual honesty. Justificationist honesty demanded the acceptance of only what was proven and the rejection of everything unproven. Neojustificationist honesty demanded the specification of the probability of any hypothesis in the light of the available empirical evidence. The honesty of naive falsificationism demanded the testing of the falsifiable and the rejection of the unfalsifiable and the falsified. Finally, the honesty of sophisticated falsificationism demanded that one should try to look at things from different points of view, to put forward new theories which anticipate novel facts, and to reject theories which have been superseded by more powerful ones.

Sophisticated methodological falsificationism blends several different traditions. From the empiricists it has inherited the determination to learn primarily from experience. From the Kantians it has taken the activist approach to the theory of knowledge. From the conventionalists it has learned the importance of decisions in methodology.

[1] Also cf. Feyerabend [1965], pp. 254–5.

[2] Popper [1959*a*], p. 87, footnote *1.

[3] Popper [1934], section 30.

[4] Cf. also *above*, p. 109, footnote 2. [*Added in press:*] Possibly it would be better in future to abandon these terms altogether, just as we have abandoned terms like 'inductive (or experimental) proof'. Then we may call (naive) 'refutations' anomalies, and (sophisticatedly) 'falsified' theories 'superseded' ones. Our 'ordinary' language is impregnated not only by 'inductivist' but also by falsificationist dogmatism. A reform is overdue.

I should like to emphasize here a further distinctive feature of sophisticated methodological empiricism: the crucial role of excess corroboration. For the inductivist, learning about a new theory is learning how much confirming evidence supports it; about refuted theories one *learns* nothing (learning, after all, is to build up proven or probable *knowledge*). For the dogmatic falsificationist, learning about a theory is learning whether it is refuted or not; about confirmed theories one learns nothing (one cannot prove or probabilify anything), about refuted theories one learns that they are disproved.[1] For the sophisticated falsificationist, learning about a theory is primarily learning which new facts it anticipated: indeed, for the sort of Popperian empiricism I advocate, the only relevant evidence is the evidence anticipated by a theory, and *empiricalness (or scientific character) and theoretical progress are inseparably connected.*[2]

This idea is not entirely new. Leibnitz, for instance, in his famous letter to Conring in 1678, wrote: 'It is the greatest commendation of an hypothesis (next to [proven] truth) if by its help predictions can be made even about phenomena or experiments not tried.'[3] Leibnitz's view was widely accepted by scientists. But since all appraisal of a scientific theory was before Popper appraisal of its degree of justification, this position was regarded by some logicians as untenable. Mill, for instance, complains in 1843 in horror that 'it seems to be thought that an hypothesis . . is entitled to a more favourable reception, if besides accounting for all the facts previously known, it has led to the anticipation and prediction of others which experience afterwards verified'.[4] Mill had a point: this appraisal was in conflict both with justificationism and with probabilism: why should an event *prove* more, if it was anticipated by the theory than if it was known already before? As long as *proof* was the only criterion of the scientific character of a theory, Leibnitz's criterion could only be regarded as irrelevant.[5] Also, the *probability* of a theory given evidence cannot possibly be influenced, as Keynes pointed out, by *when* the evidence was produced: the probability of a theory given evidence can depend only on the theory

[1] For a defence of this theory of 'learning from experience', cf. Agassi [1969].

[2] *These remarks show that 'learning from experience' is a normative idea; therefore all purely 'empirical' learning theories miss the heart of the problem.*

[3] Cf. Leibnitz [1678]. The expression in brackets shows that Leibnitz regarded this criterion as second best and thought that the best theories are those which are proved. Thus Leibnitz's position—like Whewell's—is a far cry from fully fledged sophisticated falsificationism.

[4] Mill [1843], vol. II, p. 23.

[5] This was J. S. Mill's argument (*ibid.*). He directed it against Whewell, who thought that 'consilience of inductions' or successful prediction of improbable events *verifies* (that is, *proves*) a theory. (Whewell [1858], pp. 95–6.) No doubt, *the basic contradiction both in Whewell's and in Duhem's philosophy of science is their conflation of heuristic power and proven truth. Popper separated the two.*

and the evidence,[1] and not upon whether the evidence was produced before or after the theory.

In spite of this convincing justificationist criticism, the criterion survived among some of the best scientists, since it formulated their strong dislike of merely *ad hoc* explanations, which 'though [they] truly express the facts [they set out to explain, are] not born out by any other phenomena'.[2]

But it was only Popper who recognized that the *prima facie* inconsistency between the few odd, casual remarks against *ad hoc* hypotheses on the one hand and the huge edifice of justificationist philosophy of knowledge must be solved by demolishing justificationism and by introducing new, non-justificationist criteria for appraising scientific theories based on anti-adhocness.

Let us look at a few examples. Einstein's theory is not better than Newton's *because* Newton's theory was 'refuted' but Einstein's was not: there are many known 'anomalies' to Einsteinian theory. Einstein's theory is better than—that is, represents progress compared with—Newton's theory *anno 1916* (that is, Newton's laws of dynamics, law of gravitation, the known set of initial conditions; 'minus' the list of known anomalies such as Mercury's perihelion) *because* it explained everything that Newton's theory had successfully explained, and it explained also *to some extent* some known anomalies and, in addition, forbade events like transmission of light along straight lines near large masses about which Newton's theory had said nothing but which had been permitted by other well-corroborated scientific theories of the day; moreover, *at least some* of the unexpected excess Einsteinian content was in fact *corroborated* (for instance, by the eclipse experiments).

On the other hand, according to these sophisticated standards, Galileo's theory that the natural motion of terrestrial objects was circular, introduced no improvement since it did not forbid anything that had not been forbidden by the relevant theories he intended to improve upon (that is, by Aristotelian physics and by Copernican celestial kinematics). This theory was therefore *ad hoc* and therefore—from the heuristic point of view—valueless.[3]

A beautiful example of a theory which satisfied only the first part of Popper's criterion of progress (excess content) but not the second part (corroborated excess content) was given by Popper himself: the Bohr–

[1] Keynes [1921], p. 305. But cf. my [1968*a*], p. 394.

[2] This is Whewell's critical comment on an *ad hoc* auxiliary hypothesis in Newton's theory of light (Whewell [1857], vol. II, p. 317.)

[3] In the terminology of my [1968*a*], this theory was '*ad hoc*$_1$' (cf. my [1968*a*], p. 389, footnote 1); the example was originally suggested to me by Paul Feyerabend as a paradigm of a *valuable ad hoc theory*. But cf. *below*, p. 142, especially footnote 3.

Kramers–Slater theory of 1924. This theory was refuted in *all* its new predictions.[1]

Let us finally consider how much conventionalism remains in sophisticated falsificationism. Certainly *less* than in naive falsificationism. We need *fewer* methodological decisions. The '*fourth-type decision*' which was essential for the naive version[2] has become completely redundant. To show this we only have to realize that if a scientific theory, consisting of some 'laws of nature', initial conditions, auxiliary theories (but without a *ceteris paribus* clause) conflicts with some factual propositions we do not have to decide which—explicit or 'hidden'—part to replace. We may try to replace *any* part and only when we have hit on an explanation of the anomaly with the help of some content-increasing change (or auxiliary hypothesis), and nature corroborates it, do we move on to eliminate the 'refuted' complex. Thus sophisticated falsification is a slower but possibly safer process than naive falsification.

Let us take an example. Let us assume that the course of a planet differs from the one predicted. Some conclude that this refutes the dynamics and gravitational theory applied: the initial conditions and the *ceteris paribus* clause have been ingeniously corroborated. Others conclude that this refutes the initial conditions used in the calculations: dynamics and gravitational theory have been superbly corroborated in the last two hundred years and all suggestions concerning further factors in play failed. Yet others conclude that this refutes the underlying assumption that there were no other factors in play except for those which were taken into account: these people may possibly be motivated by the metaphysical principle that any explanation is only approximative because of the infinite complexity of the factors involved in determining any single event. Should we praise the first type as '*critical*', scold the second type as '*hack*', and condemn the third as '*apologetic*'? No. We do not need to draw any conclusions about such 'refutation'. We never reject a specific theory simply by *fiat*. If we have an inconsistency like the one mentioned, we do not have to decide which ingredients of the theory we regard as problematic and which ones as unproblematic: we regard all ingredients as problematic in the light of the conflicting accepted basic statement and try to replace all of them. If we succeed in replacing some ingredient in a 'progressive' way (that is, the replacement has more corroborated empirical content than the original), we call it 'falsified'.

We do not need the *fifth type decision* of the naive falsificationist either.

[1] In the terminology of my [1968*a*], this theory was not '*ad hoc*$_1$', but it was '*ad hoc*$_2$' (cf. my [1968*a*], p. 389, footnote 1). For a simple but artificial illustration, see *ibid.* p. 387, footnote 2. (For *ad hoc*$_3$, cf. *below*, p. 175, footnote 3.)

[2] Cf. *above*, p. 110.

In order to show this let us have a new look at the problem of the appraisal of (syntactically) metaphysical theories—and the problem of their retention and elimination. The 'sophisticated' solution is obvious. We retain a syntactically metaphysical theory as long as the problematic instances can be explained by content-increasing changes in the auxiliary hypotheses appended to it.[1] Let us take, for instance, Cartesian metaphysics C: 'in *all* natural processes *there is* a clockwork mechanism regulated by (*a priori*) animating principles.' This is syntactically irrefutable: it can clash with no—spatiotemporally singular—'basic statement'. It may, of course, clash with a refutable theory like N: 'gravitation is a force equal to $fm_1 m_2/r^2$ *which acts at a distance*'. But N will only clash with C if 'action at a distance' is interpreted literally and possibly, in addition, as representing an *ultimate* truth, irreducible to any still deeper cause. (Popper would call this an 'essentialist' interpretation.) Alternatively we can regard 'action at a distance' as a mediate cause. Then we interpret 'action at a distance' figuratively, and regard it as a shorthand for some hidden mechanism of action by contact. (We may call this a 'nominalist' interpretation.) In this case we can attempt to explain N by C—Newton himself and several French physicists of the eighteenth century tried to do so. If an auxiliary theory which performs this explanation (or, if you wish, 'reduction') produces novel facts (that is, it is 'independently testable'), Cartesian metaphysics should be regarded as good, scientific, empirical metaphysics, generating a progressive problemshift. A progressive (syntactically) metaphysical theory produces a sustained progressive shift in its protective belt of auxiliary theories. If the reduction of the theory to the 'metaphysical' framework does not produce new empirical content, let alone novel facts, then the reduction represents a degenerating problemshift, it is a mere linguistic exercise. The Cartesian efforts to bolster up their 'metaphysics' in order to explain Newtonian gravitation is an outstanding example of such a merely linguistic reduction.[2]

Thus we do not eliminate a (syntactically) metaphysical theory if it clashes with a well-corroborated scientific theory, as naive falsificationism suggests. We eliminate it if it produces a degenerating shift in the long run and there is a better, rival, metaphysics to replace it. The methodology

[1] *We can formulate this condition with striking clarity only in terms of the methodology of research programmes to be explained in §3: we retain a syntactically metaphysical theory as the 'hard core' of a research programme as long as its associated positive heuristic produces a progressive problemshift in the 'protective belt' of auxiliary hypotheses.* Cf. *below*, pp. 136–7.

[2] This phenomenon was described in a beautiful paper by Whewell [1851]; but he could not explain it methodologically. Instead of recognizing the victory of the *progressive* Newtonian programme over the *degenerating* Cartesian programme, he thought this was the victory of proven truth over falsity. For details cf. my [1970]: for a general discussion of the demarcation between progressive and degenerating reduction cf. Popper [1969].

of a research programme with a 'metaphysical' core does not differ from the methodology of one with a 'refutable' core except for the logical level of the inconsistencies which are the driving force of the programme.[1]

(It has to be stressed, however, that the very choice of the logical form in which to articulate a theory depends to a large extent on our methodological decision. For instance, instead of formulating Cartesian metaphysics as an 'all-some' statement, we can formulate it as an 'all-statement': 'all natural processes are clockworks'. A 'basic statement' contradicting this would be: '*a* is a natural process and it is not clockwork'. The question is whether according to the 'experimental techniques', or rather, to the interpretative theories of the day, '*x* is not a clockwork' can be 'established' or not. Thus the rational choice of the logical form of a theory depends on the state of our knowledge; for instance, a metaphysical 'all-some' statement of today may become, with the change in the level of observational theories, a scientific 'all-statement' tomorrow. I have already argued that only series of theories and not theories should be classified as scientific or non-scientific; now I have indicated that even the logical form of a theory can only be rationally chosen on the basis of a critical appraisal of the state of the research programme in which it is embedded.)

The first, second, and third type decisions of naive falsificationism[2] however cannot be avoided, but as we shall show, the conventional element in the second decision—and also in the third—can be slightly reduced. We cannot avoid the decision which sort of propositions should be the 'observational' ones and which the 'theoretical' ones. We cannot avoid either the decision about the truth-value of some 'observational propositions'. These decisions are vital for the decision whether a problemshift is empirically progressive or degenerating.[3] But the sophisticated falsificationist may at least mitigate the arbitrariness of this second decision by allowing for an *appeal procedure.*

Naive falsificationists do not lay down any such appeal procedure. They accept a basic statement if it is backed up by a well-corroborated falsifying hypothesis,[4] and let it overrule the theory under test—even though they are well aware of the risk.[5] But there is no reason why we should not regard a falsifying hypothesis—and the basic statement it supports—as being just as problematic as a falsified hypothesis. Now how exactly can we expose the problematicality of a basic statement? On what grounds can the protagonists of the 'falsified' theory appeal and win?

Some people may say that we might go on testing the basic statement (or

[1] Cf. *above*, p. 126, footnote 1.
[2] Cf. *above*, pp. 106 and 109.
[3] Cf. *above*, p. 118.
[4] Popper [1934], section 22.
[5] Cf. e.g. Popper [1959*a*], p. 107, footnote *2. Also cf. *above*, pp. 112–14.

the falsifying hypothesis) 'by their deductive consequences' until agreement is finally reached. In this testing we deduce—in the same deductive model—further consequences from the basic statement either with the help of the theory under test or some other theory which we regard as unproblematic. Although this procedure 'has no natural end', we always come to a point when there is no further disagreement.[1]

But when the theoretician appeals against the verdict of the experimentalist, the appeal court does not normally cross-question the basic statement directly but rather questions the *interpretative theory* in the light of which its truth-value had been established.

One typical example of a series of successful appeals is the Proutians' fight against unfavourable experimental evidence from 1815 to 1911. For decades Prout's theory T ('that all atoms are compounds of hydrogen atoms and thus "atomic weights" of all chemical elements must be expressible as whole numbers') and falsifying 'observational' hypotheses, like Stas's 'refutation' R ('the atomic weight of chlorine is 35·5') confronted each other. As we know, in the end T prevailed over R.[2]

The first stage of any serious criticism of a scientific theory is to reconstruct, improve, its logical deductive articulation. Let us do this in the case of Prout's theory *vis à vis* Stas's refutation. First of all, we have to realize that in the formulation we just quoted, T and R were *not* inconsistent. (Physicists rarely articulate their theories sufficiently to be pinned down and caught by the critic.) In order to show them up as inconsistent we have to put them in the following form. T: 'the atomic weight of all pure (homogeneous) chemical elements are multiples of the atomic weight of hydrogen', and R: 'chlorine is a pure (homogeneous) chemical element and its atomic weight is 35·5'. The last statement is in the form of a falsifying hypothesis which, if well corroborated, would allow us to use basic statements of the form B: 'Chlorine X is a pure (homogeneous) chemical element and its atomic weight is 35·5'—where X is the proper name of a 'piece' of chlorine determined, say, by its space-time co-ordinates.

But how well-corroborated is R? The first component of it says that R_1: 'Chlorine X is a pure chemical element.' This was the verdict of the experimental chemist after a rigorous application of the 'experimental techniques' of the day.

Let us have a closer look at the fine-structure of R_1. In fact R_1 stands for a conjunction of two longer statements T_1 and T_2. The first statement,

[1] This is argued in Popper [1934], section 29.

[2] Agassi claims that this example shows that we may 'stick to the hypothesis in the face of known facts in the hope that the facts will adjust themselves to theory rather than the other way round' ([1966], p. 18). But *how* can facts 'adjust themselves'? Under which *particular* conditions should the theory win? Agassi gives no answer.

T_1, could be this: 'If seventeen chemical purifying procedures $p_1, p_2 \ldots p_{17}$ are applied to a gas, what remains will be pure chlorine.' T_2 is then: 'X was subjected to the seventeen procedures $p_1, p_2 \ldots p_{17}$.' The careful 'experimenter' carefully applied all seventeen procedures: T_2 is to be accepted. But the conclusion that therefore what remained *must* be pure chlorine is a 'hard fact' only in virtue of T_1. The experimentalist, while *testing* T, *applied* T_1. He *interpreted* what he saw in the light of T_1: the result was R_1. *Yet in the monotheoretical model of the explanatory theory under test this interpretative theory does not appear at all.*

But what if T_1, the interpretative theory, is false? Why not 'apply' T rather than T_1 and claim that atomic weights *must be* whole numbers? Then *this* will be a 'hard fact' in the light of T, and T_1 will be overthrown. Perhaps additional new purifying procedures must be invented and applied.

The problem is then *not* when we should stick to a *'theory'* in the face of *'known facts'* and when the other way round. The problem is *not* what to do when 'theories' clash with 'facts'. Such a 'clash' is only suggested by the *'monotheoretical deductive model'*. Whether a proposition is a *'fact'* or a *'theory'* in the context of a test-situation depends on our methodological decision. 'Empirical basis of a theory' is a mono-theoretical notion, it is *relative* to some mono-theoretical deductive structure. We may use it as first approximation; but in case of 'appeal' by the theoretician, we must use a *pluralistic model.* In the pluralistic model the clash is not 'between theories and facts' but between two high-level theories: between an *interpretative theory* to provide the facts and an *explanatory theory* to explain them; and the interpretative theory may be on quite as high a level as the explanatory theory. The clash is then not any more between a logically higher-level theory and a lower-level falsifying hypothesis. The problem should not be put in terms of whether a *'refutation'* is real or not. The problem is how to repair an *inconsistency* between the 'explanatory theory' under test and the —explicit or hidden—'interpretative' theories; or, if you wish, *the problem is which theory to consider as the interpretative one which provides the 'hard' facts and which the explanatory one which 'tentatively' explains them.* In a mono-theoretical model we regard the higher-level theory as an *explanatory theory to be judged by the 'facts'* delivered from outside (by the authoritative experimentalist): in the case of a clash we reject the explanation.[1] In a pluralistic model we may decide, alternatively, to regard the higher-

[1] The decision to use some monotheoretical model is clearly vital for the naive falsificationist to enable him to reject a theory on the *sole* ground of experimental evidence. *It is in line with the necessity for him to divide sharply, at least in a test-situation, the body of science into two: the problematic and the unproblematic.* (Cf. *above* p. 107.) *It is only the theory he decides to regard as problematic which he articulates in his deductive model of criticism.*

level theory as an *interpretative theory to judge the 'facts'* delivered from outside: in case of a clash we may reject the 'facts' as 'monsters'. In a pluralistic model of testing, several theories—more or less deductively organized—are soldered together.

This argument alone would be enough to show the correctness of the conclusion, which we drew from a different earlier argument, that experiments do not simply overthrow theories, that no theory forbids a state of affairs specifiable in advance.[1] It is not that we propose a theory and Nature may shout NO; rather, we propose a maze of theories, and Nature may shout INCONSISTENT.[2]

The problem is then *shifted* from the old problem of replacing a theory refuted by 'facts' to the new problem of how to resolve inconsistencies between closely associated theories. Which of the mutually inconsistent theories should be eliminated? The sophisticated falsificationist can answer that question easily: one had to try to replace first one, then the other, then possibly both, and opt for that new set-up which provides the biggest increase in corroborated content, which provides the most progressive problemshift.[3]

Thus we have established an appeal procedure in case the theoretician wishes to question the negative verdict of the experimentalist. The theoretician may demand that the experimentalist specify his 'interpretative theory',[4] and he may then replace it—to the experimentalist's annoyance—by a better one in the light of which his originally 'refuted' theory may receive positive appraisal.[5]

[1] Cf. *above*, p. 100.

[2] Let me here answer a possible objection: 'Surely we do not need Nature to tell us that a set of theories is *inconsistent*. Inconsistency—unlike falsehood—can be ascertained without Nature's help'. But Nature's actual 'NO' in a monotheoretical methodology takes the form of a fortified 'potential falsifier', that is a sentence which, in this way of speech, we claim Nature had uttered and which is the *negation of our theory*. Nature's actual 'INCONSISTENCY' in a pluralistic methodology takes the form of a 'factual' statement couched in the light of one of the theories involved, which we claim Nature had uttered and which, if added to our proposed theories, yields an *inconsistent system*.

[3] For instance, in our earlier example (cf. *above*, p. 107 ff.) some may try to replace the gravitational theory with a new one and others may try to replace the radio-optics by a new one: we choose the way which offers the more spectacular growth, the more progressive problemshift.

[4] Criticism does not *assume* a fully articulated deductive structure: it creates it. (Incidentally, this is the main message of my [1963–4].)

[5] A classical example of this pattern is Newton's relation to Flamsteed, the first Astronomer Royal. For instance, Newton visited Flamsteed on 1 September 1694, when working full time on his lunar theory; told him to reinterpret some of his data since they contradicted his own theory; and he explained to him exactly how to do it. Flamsteed obeyed Newton and wrote to him on 7 October: 'Since you went home, I examined the observations I employed for determining the greatest equations of the earth's orbit, and considering the moon's places at the times . . . ' I find that (*if, as you intimate, the earth inclines*

But even this appeal procedure cannot do more than *postpone* the conventional decision. For the verdict of the appeal court is not infallible either. When we decide whether it is the replacement of the 'interpretative' or of the 'explanatory' theory that produces novel facts, we again must take a decision about the acceptance or rejection of basic statements. But then we have only *postponed*—and possibly *improved*—the decision, not avoided it.[1] The difficulties concerning the empirical basis which confronted 'naive' falsificationism cannot be avoided by 'sophisticated' falsificationism either. Even if we regard a theory as 'factual', that is, if our slow-moving and limited imagination cannot offer an alternative to it (as Feyerabend used to put it), we have to make, at least occasionally and temporarily, decisions about its truth-value. *Even then, experience still remains, in an important sense, the 'impartial arbiter'[2] of scientific controversy.* We cannot get rid of the problem of the 'empirical basis', if we want to learn from experience[3]: but we can make our learning less dogmatic—but also less fast and less dramatic. By regarding some observational theories as problematic we may make our methodology more flexible: but we cannot articulate and include *all* 'background knowledge' (or 'background ignorance'?) into our critical deductive model. This process is bound to be piecemeal and some conventional line must be drawn at any given time.

There is one objection even to the sophisticated version of methodological falsificationism which cannot be answered without some concession to Duhemian 'simplicism'. The objection is the so-called 'tacking paradox'. According to our definitions, adding to a theory completely disconnected low-level hypotheses may constitute a 'progressive shift'. It is difficult to eliminate such makeshift shifts without demanding that 'the additional assertions must be connected with the contradicting assertion *more intimately* than by mere conjunction'.[4] This, of course, is a sort of simplicity

on that side the moon then is) you may abate abt 20″ from it . . .' Thus Newton constantly criticized and corrected Flamsteed's observational theories. Newton taught Flamsteed, for instance, a better theory of the refractive power of the atmosphere; Flamsteed accepted this and corrected his original 'data'. One can understand the constant humiliation and slowly increasing fury of this great observer, having his data criticized and improved by a man who, on his own confession, made no observations himself: it was this feeling—I suspect—which led finally to a vicious personal controversy.

[1] The same applies to the third type of decision. If we reject a stochastic hypothesis only for one which, in our sense, supersedes it, the exact form of the 'rejection rules' becomes *less* important.

[2] Popper [1945], vol. II, chapter 23, p. 218.

[3] Agassi is then wrong in his thesis that 'observation reports may be accepted as false and hence the problem of the empirical basis is thereby disposed of' (Agassi [1966], p. 20).

[4] Feyerabend [1965], p. 226.

requirement which would assure the continuity in the series of theories which can be said to constitute *one* problemshift.

This leads us to further problems. For one of the crucial features of sophisticated falsificationism is that it replaces the concept of *theory* as the basic concept of the logic of discovery by the concept of *series of theories. It is a succession of theories and not one given theory which is appraised as scientific or pseudo-scientific.* But the members of such series of theories are usually connected by a remarkable *continuity* which welds them into *research programmes.* This *continuity*—reminiscent of Kuhnian 'normal science'—plays a vital role in the history of science; the main problems of the logic of discovery cannot be satisfactorily discussed except in the framework of a *methodology of research programmes.*

3. A METHODOLOGY OF SCIENTIFIC RESEARCH PROGRAMMES

I have discussed the problem of objective appraisal of scientific growth in terms of progressive and degenerating problemshifts in series of scientific theories. The most important such series in the growth of science are characterized by a certain *continuity* which connects their members. This continuity evolves from a genuine research programme adumbrated at the start.[1] The programme consists of methodological rules: some tell us what paths of research to avoid (*negative heuristic*), and others what paths to pursue (*positive heuristic*).

Even science as a whole can be regarded as a huge research programme with Popper's supreme heuristic rule: 'devise conjectures which have more empirical content than their predecessors.' Such methodological rules may be formulated, as Popper pointed out, as metaphysical principles.[2] For instance, the *universal* anti-conventionalist rule against exception-barring may be stated as the metaphysical principle: 'Nature does not allow exceptions'. This is why Watkins called such rules 'influential metaphysics'.[3]

But what I have primarily in mind is not science as a whole, but rather *particular* research programmes, such as the one known as 'Cartesian metaphysics'. Cartesian metaphysics, that is, the mechanistic theory of the

[1] One may point out that the negative and positive heuristic gives a rough (implicit) definition of the 'conceptual framework' (and consequently of the language). The recognition that the history of science is the history of research programmes rather than of theories may therefore be seen as a partial vindication of the view that the history of science is the history of conceptual frameworks or of scientific languages.

[2] Popper [1934], sections 11 and 70. I use 'metaphysical' as a technical term of naive falsificationism: a contingent proposition is 'metaphysical' if it has no 'potential falsifiers'.

[3] Watkins [1958]. Watkins cautions that 'the logical gap between statements and prescriptions in the metaphysical-methodological field is illustrated by the fact that a person may reject a [metaphysical] doctrine in its fact-stating form while subscribing to the prescriptive version of it' (*Ibid.* pp. 356–7).

universe—according to which the universe is a huge clockwork (and system of vortices) with push as the only cause of motion—functioned as a powerful heuristic principle. It discouraged work on scientific theories—like [the 'essentialist' version of] Newton's theory of action at a distance—which were inconsistent with it (*negative heuristic*). On the other hand, it encouraged work on auxiliary hypotheses which might have saved it from apparent counterevidence—like Keplerian ellipses (*positive heuristic*).[1]

(*a*) *Negative heuristic: the 'hard core' of the programme.*

All scientific research programmes may be characterized by their *'hard core'*. The negative heuristic of the programme forbids us to direct the *modus tollens* at this 'hard core'. Instead, we must use our ingenuity to articulate or even invent 'auxiliary hypotheses', which form a *protective belt* around this core, and we must redirect the *modus tollens* to *these.* It is this protective belt of auxiliary hypotheses which has to bear the brunt of tests and get adjusted and re-adjusted, or even completely replaced, to defend the thus-hardened core. A research programme is successful if all this leads to a progressive problemshift; unsuccessful if it leads to a degenerating problemshift.

The classical example of a successful research programme is Newton's gravitational theory: possibly the most successful research programme ever. When it was first produced, it was submerged in an ocean of 'anomalies' (or, if you wish, 'counterexamples'[2]), and opposed by the observational theories supporting these anomalies. But Newtonians turned, with brilliant tenacity and ingenuity, one counter-instance after another into corroborating instances, primarily by overthrowing the original observational theories in the light of which this 'contrary evidence' was established. In the process they themselves produced new counter-examples which they again resolved. They 'turned each new difficulty into a new victory of their programme'.[3]

In Newton's programme the negative heuristic bids us to divert the *modus tollens* from Newton's three laws of dynamics and his law of gravitation. This 'core' is 'irrefutable' by the methodological decision of its protagonists: anomalies must lead to changes only in the 'protective' belt of auxiliary, 'observational' hypothesis and initial conditions.[4]

I have given a contrived micro-example of a progressive Newtonian

[1] For this Cartesian research programme, cf. Popper [1958] and Watkins [1958], pp. 350–1.

[2] For the clarification of the concepts of 'counterexample' and 'anomaly' cf. *above*, p. 110, and especially *below*, p. 159, footnote 1.

[3] Laplace [1796], livre IV, chapter ii.

[4] The actual hard core of a programme does not actually emerge fully armed like Athene from the head of Zeus. It develops slowly, by a long, preliminary process of trial and error. In this paper this process is not discussed.

problemshift.[1] If we analyse it, it turns out that each successive link in this exercise predicts some new fact; each step represents an increase in empirical content: the example constitutes a *consistently progressive theoretical shift*. Also, each prediction is in the end verified; although on three subsequent occasions they may have seemed momentarily to be 'refuted'.[2] While 'theoretical progress' (in the sense here described) may be verified immediately,[3] 'empirical progress' cannot, and in a research programme we may be frustrated by a long series of 'refutations' before ingenious and lucky content-increasing auxiliary hypotheses turn a chain of defeats—*with hindsight*—into a resounding success story, either by revising some false 'facts' or by adding novel auxiliary hypotheses. We may then say that we must require that each step of a research programme be consistently content-increasing: that each step constitute a *consistently progressive theoretical problemshift*. All we need in addition to this is that at least every now and then the increase in content should be seen to be retrospectively corroborated: the programme as a whole should also display an *intermittently progressive empirical shift*. We do not demand that each step produce *immediately* an *observed* new fact. Our term '*intermittently*' gives sufficient *rational* scope for dogmatic adherence to a programme in face of *prima facie* 'refutations'.

The idea of 'negative heuristic' of a scientific research programme rationalizes classical conventionalism to a considerable extent. We may rationally decide not to allow 'refutations' to transmit falsity to the hard core as long as the corroborated empirical content of the protecting belt of auxiliary hypotheses increases. But our approach differs from Poincaré's justificationist conventionalism in the sense that, unlike Poincaré's, we maintain that if and when the programme ceases to anticipate novel facts, its hard core might have to be abandoned: that is, *our* hard core, unlike Poincaré's, may crumble under certain conditions. In this sense we side with Duhem who thought that such a possibility must be allowed for[4]; but for Duhem the reason for such crumbling is purely *aesthetic*,[5] while for us it is mainly *logical and empirical*.

(*b*) *Positive heuristic: the construction of the 'protective belt' and the relative autonomy of theoretical science.*

Research programmes, besides their negative heuristic, are also characterized by their positive heuristic.

[1] Cf. *above*, pp. 100–1. For *real* examples, cf. my [1970].

[2] The 'refutation' was each time successfully diverted to 'hidden lemmas'; that is, to lemmas emerging, as it were, from the *ceteris paribus* clause.

[3] But cf. *below*, pp. 155–7. [4] Cf. *above*, p. 105. [5] *Ibid.*

Even the most rapidly and consistently progressive research programmes can digest their 'counter-evidence' only piecemeal: anomalies are never completely exhausted. But it should not be thought that yet unexplained anomalies—'puzzles' as Kuhn might call them—are taken in random order, and the protective belt built up in an eclectic fashion, without any preconceived order. The order is usually decided in the theoretician's cabinet, independently of the *known* anomalies. Few theoretical scientists engaged in a research programme pay undue attention to 'refutations'. They have a long-term research policy which anticipates these refutations. This research policy, or order of research, is set out—in more or less detail—in the *positive heuristic* of the research programme. The negative heuristic specifies the 'hard core' of the programme which is 'irrefutable' by the methodological decision of its protagonists; the positive heuristic consists of a partially articulated set of suggestions or hints on how to change, develop the 'refutable variants' of the research-programme, how to modify, sophisticate, the 'refutable' protective belt.

The positive heuristic of the programme saves the scientist from becoming confused by the ocean of anomalies. The positive heuristic sets out a programme which lists a chain of ever more complicated *models* simulating reality: the scientist's attention is riveted on building his models following instructions which are laid down in the positive part of his programme. He ignores the *actual* counterexamples, the available '*data*'.[1] Newton first worked out his programme for a planetary system with a fixed point-like sun and one single point-like planet. It was in this model that he derived his inverse square law for Kepler's ellipse. But this model was forbidden by Newton's own third law of dynamics, therefore the model had to be replaced by one in which both sun and planet revolved round their common centre of gravity. This change was not motivated by any observation (the data did not suggest an 'anomaly' here) but by a theoretical difficulty in developing the programme. Then he worked out the programme for more planets as if there were only heliocentric but no interplanetary forces. Then he worked out the case where the sun and planets were not mass-points but mass-*balls*. Again, for this change he did not *need* the observation of an anomaly; infinite density was forbidden by an (inarticulated) touchstone theory, therefore planets *had* to be extended. This change involved considerable mathematical difficulties, held up Newton's work—and delayed the publication of the *Principia* by more than a decade. Having solved this 'puzzle',

[1] If a scientist (or mathematician) has a positive heuristic, he refuses to be drawn into observation. He will 'lie down on his couch, shut his eyes and forget about the data'. (Cf. my [1963–4], especially pp. 300 ff., where there is a detailed case study of such a programme.) Occasionally, of course, he will ask Nature a shrewd question: he will then be encouraged by Nature's *YES*, but not discouraged by its *NO*.

he started work on *spinning balls* and their wobbles. Then he admitted interplanetary forces and started work on *perturbations*. At this point he started to look more anxiously at the facts. Many of them were beautifully explained (qualitatively) by this model, many were not. It was then that he started to work on *bulging* planets, rather than round planets, etc.

Newton despised people who, like Hooke, stumbled on a first naive model but did not have the tenacity and ability to develop it into a research programme, and who thought that a first version, a mere aside, constituted a 'discovery'. He held up publication until his programme had achieved a remarkable progressive shift.[1]

Most, if not all, Newtonian 'puzzles', leading to a series of new variants superseding each other, were forseeable at the time of Newton's first naive model and no doubt Newton and his colleagues *did* forsee them: Newton must have been fully aware of the blatant falsity of his first variants.[2] Nothing shows the existence of a positive heuristic of a research programme clearer than this fact: this is why one speaks of 'models' in research programmes. A '*model*' is a set of initial conditions (possibly together with some of the observational theories) which one knows is *bound* to be replaced during the further development of the programme, and one even knows, more or less, how. This shows once more how irrelevant 'refutations' of any specific variant are in a research programme: their existence is fully expected, the positive heuristic is there as the strategy both for predicting (producing) and digesting them. Indeed, if the positive heuristic is clearly spelt out, the difficulties of the programme are mathematical rather than empirical.[3]

One may formulate the 'positive heuristic' of a research programme as a 'metaphysical' principle. For instance one may formulate Newton's programme like this: 'the planets are essentially gravitating spinning-tops of roughly spherical shape'. This idea was never *rigidly* maintained: the planets are not *just* gravitational, they have also, for example, electromagnetic characteristics which may influence their motion. Positive heuristic is thus in general more flexible than negative heuristic. Moreover, it occasionally happens that when a research programme gets into a degenerating phase, a little revolution or a *creative shift* in its positive heuristic may push it forward again.[1] It is better therefore to separate the 'hard core' from the more flexible metaphysical principles expressing the positive heuristic.

Our considerations show that the positive heuristic forges ahead with almost complete disregard of 'refutations': it may seem that it is the '*verifications*'[2] rather than the refutations which provide the contact points with reality. Although one must point out that any 'verification' of the $n+1$-th version of the programme is a refutation of the n-th version, we cannot deny that *some* defeats of the subsequent versions are always foreseen: it is the 'verifications' which keep the programme going, recalcitrant instances notwithstanding.

We may appraise research programmes, even after their 'elimination', for their *heuristic power*: how many new facts did they produce, how great was 'their capacity to explain their refutations in the course of their growth'?[3]

(We may also appraise them for the stimulus they gave to mathematics. The real difficulties for the theoretical scientist arise rather from the *mathematical difficulties* of the programme than from anomalies. The greatness of the Newtonian programme comes partly from the development—by Newtonians—of classical infinitesimal analysis which was a crucial precondition of its success.)

Thus the methodology of scientific research programmes accounts for the *relative autonomy of theoretical science:* a historical fact whose rationality cannot be explained by the earlier falsificationists. Which problems scientists working in powerful research programmes rationally choose, is determined by the positive heuristic of the programme rather than by psychologically worrying (or technologically urgent) anomalies. The anomalies are listed but shoved aside in the hope that they will turn, in due course, into corroborations of the programme. Only those scientists have to rivet their attention on anomalies who are either engaged in trial-and-error exercises[4] or who work in a degenerating phase of a research programme when the positive heuristic ran out of steam. (All this, of course, must sound repugnant to naive falsificationists who hold that once

[1] Reichenbach, following Cajori, gives a different explanation of what delayed Newton in the publication of his *Principia*: 'To his disappointment he found that the observational results disagreed with his calculations. Rather than set any theory, however beautiful, before the facts, Newton put the manuscript of his theory into his drawer. Some twenty years later, after new measurements of the circumference of the earth had been made by a French expedition, Newton saw that the figures on which he had based his test were false and that the improved figures agreed with his theoretical calculation. It was only after this test that he published his law . . . The story of Newton is one of the most striking illustrations of the method of modern science' (Reichenbach [1951], pp. 101–2). Feyerabend criticizes Reichenbach's account (Feyerabend [1965], p. 229), but does not give an alternative *rationale*.

[2] For a further discussion of Newton's research programme, cf. my [1970].

[3] For this point cf. Truesdell [1960].

[1] Soddy's contribution to Prout's programme or Pauli's to Bohr's (old quantum theory) programme are typical examples of such creative shifts.

[2] A 'verification' is a corroboration of excess content in the expanding programme. But, of course, a 'verification' does not *verify* a programme: it shows only its heuristic power.

[3] Cf. my [1963–4], pp. 324–30. Unfortunately in 1963–4 I had not yet made a clear terminological distinction between theories and research programmes, and this impaired my exposition of a research programme in informal, quasi-empirical mathematics. There are fewer such shortcomings in my [1971].

[4] Cf. *below*, p. 175.

a theory is 'refuted' by experiment (by *their* rule book), it is irrational (and dishonest) to develop it further: one has to replace the old 'refuted' theory by a new, unrefuted one.)

(*c*) *Two illustrations: Prout and Bohr.*

The dialectic of positive and negative heuristic in a research programme can best be illuminated by examples. Therefore I am now going to sketch a few aspects of two spectacularly successful research programmes: Prout's programme[1] based on the idea that all atoms are compounded of hydrogen atoms and Bohr's programme based on the idea that light-emission is due to electrons jumping from one orbit to another within the atoms.

(*In writing a historical case study, one should, I think, adopt the following procedure:* (1) *one gives a rational reconstruction;* (2) *one tries to compare this rational reconstruction with actual history and to criticize both one's rational reconstruction for lack of historicity and the actual history for lack of rationality. Thus any historical study must be preceded by a heuristic study: history of science without philosophy of science is blind. In this paper it is not my purpose to go on seriously to the second stage.*)

(*c* 1) *Prout: a research programme progressing in an ocean of anomalies.*

Prout, in an anonymous paper of 1815, claimed that the atomic weights of all pure chemical elements were whole numbers. He knew very well that anomalies abounded, but said that these arose because chemical substances as they ordinarily occurred were *impure*: that is, the relevant 'experimental techniques' of the time were unreliable, or, to put it in our terms, the contemporary 'observational' theories in the light of which the truth values of the basic statements of his theory were established, were false.[2] The champions of Prout's theory therefore embarked on a major venture: to overthrow those theories which supplied the counter-evidence to their thesis. For this they had to revolutionize the established analytical chemistry of the time and correspondingly revise the experimental techniques with which pure elements were to be separated.[3]

[1] Already mentioned *above*, pp. 128–9.

[2] Alas, all this is rational reconstruction rather than actual history. Prout denied the existence of any anomalies. For instance, he claimed that the atomic weight of chlorine was exactly 36.

[3] Prout was aware of some of the basic methodological features of his programme. Let us quote the first lines of his [1815]: 'The author of the following essay submits it to the public with the greatest diffidence... He trusts, however, that its importance will be seen, and that some one will undertake to examine it, and thus verify or refute its conclusions. If these should be proved erroneous, still new facts may be brought to light, or old ones better established, by the investigation; but if they should be verified, a new and interesting light will be thrown upon the whole science of chemistry.'

Prout's theory, as a matter of fact, defeated the theories previously applied in purification of chemical substances one after the other. Even so, some chemists became tired of the research programme and gave it up, since the successes were still far from adding up to a final victory. For instance, Stas, frustrated by some stubborn, recalcitrant instances, concluded in 1860 that Prout's theory was 'without foundations'.[1] But others were more encouraged by the progress than discouraged by the lack of complete success. For instance, Marignac immediately retorted that 'although [he is satisfied that] the experiments of Monsieur Stas are perfectly exact, [there is no proof] that the differences observed between his results and those required by Prout's law cannot be explained by the imperfect character of experimental methods'.[2] As Crookes put it in 1886: 'Not a few chemists of admitted eminence consider that we have here [in Prout's theory] an expression of the truth, masked by some residual or collateral phenomena which we have not yet succeeded in eliminating.'[3] That is, there had to be some *further* false hidden assumption in the 'observational' theories on which 'experimental techniques' for chemical purification were based and with the help of which atomic weights were calculated: in Crookes's view even in 1886 'some present atomic weights merely represented a mean value'.[4] Indeed, Crookes went on to put this idea in a scientific (content-increasing) form: he proposed concrete new theories of 'fractionation', a new 'sorting Demon'.[5] But, alas, his new observational theories turned out to be as false as they were bold and, being unable to anticipate any new fact, they were eliminated from the (rationally reconstructed) history of science. As it turned out a generation later, there was a very basic hidden assumption which failed the researchers: that two pure elements must be separable by *chemical* methods. The idea that two different pure elements may behave identically in all *chemical* reactions but can be separated by *physical* methods, required a change, a '*stretching*', of the concept of 'pure element' which constituted a change—a *concept-stretching expansion*—of the research programme itself.[6] This revolutionary highly *creative shift* was taken only by Rutherford's school[7]; and then 'after

[1] Clerk Maxwell was on Stas's side: he thought it was impossible that there should be two kinds of hydrogen, 'for if some [molecules] were of slightly greater mass than others, we have the means of producing a separation between molecules of different masses, one of which would be somewhat denser than the other. As this cannot be done, we must admit [that all are alike]' (Maxwell [1871]). [2] Marignac [1860].

[3] Crookes [1886]. [4] *Ibid.* [5] Crookes [1886], p. 491.

[6] For 'concept-stretching', cf. my [1963–4], part IV.

[7] The shift is anticipated in Crookes's fascinating [1888] where he indicates that the solution should be sought in a new demarcation between 'physical' and 'chemical'. But the anticipation remained philosophical; it was left to Rutherford and Soddy to develop it, after 1910, into a scientific theory.

many vicissitudes and the most convincing apparent disproofs, the hypothesis thrown out so lightly by Prout, an Edinburgh physician, in 1815, has, a century later, become the corner-stone of modern theories of the structure of atoms'.[1] However, this creative step was in fact only a side-result of progress in a different, indeed, distant research programme; Proutians, lacking this *external* stimulus, never dreamt of trying, for instance, to build powerful centrifugal machines to separate elements.

(When an 'observational' or 'interpretative' theory finally gets eliminated, the 'precise' measurements carried out within the discarded framework may look—with hindsight—rather foolish. Soddy made fun of 'experimental precision' for its own sake: 'There is something surely akin to if not transcending tragedy in the fate that has overtaken the life work of that distinguished galaxy of nineteenth-century chemists, rightly revered by their contemporaries as representing the crown and perfection of accurate scientific measurement. Their hard won results, for the moment at least, appears as of as little interest and significance as the determination of the average weight of a collection of bottles, some of them full and some of them more or less empty.'[2])

Let us stress that in the light of the methodology of research programmes here proposed there never was any rational reason to *eliminate* Prout's programme. Indeed, the programme produced a beautiful, progressive shift, even if, in between, there were considerable hitches.[3] Our sketch shows how a research programme can challenge a considerable bulk of accepted scientific knowledge: it is planted, as it were, in an inimical environment which, step by step, it can override and transform.

Also, the actual history of Prout's programme illustrates only too well how much the progress of science was hindered and slowed down by justificationism and by naive falsificationism. (The opposition to atomic theory in the nineteenth century was fostered by both.) An elaboration of this particular influence of bad methodology on science may be a rewarding research programme for the historian of science.

(*c* 2) *Bohr: a research programme progressing on inconsistent foundations.*

A brief sketch of Bohr's research programme of light emission (in *early* quantum physics) will illustrate further—and even expand—our thesis.[4]

[1] Soddy [1932], p. 50. [2] *Ibid.*

[3] These hitches inevitably induce many individual scientists to shelve or altogether jettison the programme and join other research programmes where the positive heuristic happens to offer at the time cheaper successes: the history of science cannot be *fully* understood without mob-psychology. (Cf. *below*, pp. 177–80.)

[4] This section may again strike the historian as more a caricature than a sketch; but I hope it serves its purpose. (Cf. *above*, p. 138.) Some statements are to be taken not with a grain, but with tons, of salt.

The story of Bohr's research programme can be characterized by: (1) its initial problem; (2) its negative and positive heuristic; (3) the problems which it attempted to solve in the course of its development; and (4) its degeneration point (or, if you wish, 'saturation point') and, finally, (5) the programme by which it was superseded.

The background problem was the riddle of how Rutherford atoms (that is, minute planetary systems with electrons orbiting round a positive nucleus) can remain stable; for, according to the well-corroborated Maxwell–Lorentz theory of electromagnetism they should collapse. But Rutherford's theory was well corroborated too. Bohr's suggestion was to ignore for the time being the inconsistency and consciously develop a research programme whose 'refutable' versions were inconsistent with the Maxwell–Lorentz theory.[1] He proposed five postulates as the *hard core* of his programme: '(1) that energy radiation [within the atom] is not emitted (or absorbed) in the continuous way assumed in the ordinary electrodynamics, but only during the passing of the systems between different "stationary" states. (2) That the dynamical equilibrium of the systems in the stationary states is governed by the ordinary laws of mechanics, while these laws do not hold for the passing of the systems between the different states. (3) That the radiation emitted during the transition of a system between two stationary states is homogeneous, and that the relation between the frequency ν and the total amount of energy emitted E is given by $E = h\nu$, where h is Planck's constant. (4) That the different stationary states of a simple system consisting of an electron rotating round a positive nucleus are determined by the condition that the ratio between the total energy, emitted during the formation of the configuration, and the frequency of revolution of the electron is an entire multiple of $\frac{1}{2}h$. Assuming that the orbit of the electron is circular, this assumption is equivalent with the assumption that the angular momentum of the electron round the nucleus is equal to an entire multiple of $h/2\pi$. (5) That the "permanent" state of any atomic system, i.e. the state in which the energy emitted is maximum, is determined by the condition that the angular momentum of every electron round the centre of its orbit is equal to $h/2\pi$.'[2]

We have to appreciate the crucial methodological difference between the inconsistency introduced by Prout's programme and that introduced by Bohr's. Prout's research programme declared war on the analytical chemistry of his time: its positive heuristic was designed to overthrow it and replace it. But Bohr's research programme contained no analogous design:

[1] This, of course, is a further argument against J. O. Wisdom's thesis that metaphysical theories can be refuted by a conflicting well corroborated scientific theory (Wisdom [1963]) Also, cf. *above*, p. 112, text to footnote 1, and pp. 126–7.

[2] Bohr [1913*a*], p. 874.

its positive heuristic, even if it had been completely successful, would have left the inconsistency with the Maxwell–Lorentz theory unresolved.[1] To suggest such an idea required even greater courage than Prout's; the idea crossed Einstein's mind but he found it unacceptable, and rejected it.[2] Indeed, *some of the most important research programmes in the history of science were grafted on to older programmes with which they were blatantly inconsistent.* For instance, Copernican astronomy was 'grafted' on to Aristotelian physics, Bohr's programme on to Maxwell's. Such 'grafts' are irrational for the justificationist and for the naive falsificationist, neither of whom can countenance growth on inconsistent foundations. Therefore they are usually concealed by *ad hoc* stratagems—like Galileo's theory of circular inertia or Bohr's correspondence, and, later, complementarity principle—the only purpose of which is to hide the 'deficiency'.[3] As the young grafted programme strengthens, the peaceful co-existence comes to an end, the symbiosis becomes competitive and the champions of the new programme try to replace the old programme altogether.

It may well have been the success of his 'grafted programme' which later misled Bohr into believing that such fundamental inconsistencies in research programmes can and should be put up with *in principle*, that they do not present any serious problem and one merely has to get used to them. Bohr tried in 1922 to lower the standards of scientific criticism; he argued that '*the most* that one can demand of a theory [i.e. programme] is that the classification [it establishes] can be pushed so far that it can contribute to the development of the field of observation by the prediction of *new* phenomena.'[4]

(This statement by Bohr is similar to d'Alembert's when faced with the inconsistency in the foundations of infinitesimal theory: '*Allez en avant et la foi vous viendra.*' According to Margenau, 'it is understandable that, in the excitement over its success, men overlooked a malformation in the theory's architecture; for Bohr's atom sat like a baroque tower upon the Gothic base of classical electrodynamics.'[5] But as a matter of fact, the 'malformation' was not 'overlooked': everybody was aware of it, only they ignored it—more or less—during the progressive phase of the programme.[6]

[1] Bohr held at this time that the Maxwell–Lorentz theory would *eventually* have to be replaced (Einstein's photon theory had already indicated this need).

[2] Hevesy [1913]; cf. also *above*, p. 136, text to footnote 1.

[3] In our methodology there is no need for such protective *ad hoc* stratagems. But, on the other hand, they are harmless as long as they are clearly seen as problems, not as solutions.

[4] Bohr [1922]; my italics.

[5] Margenau [1950], p. 311.

[6] Sommerfeld ignored it more than Bohr: cf. *below*, p. 150, footnote 4.

Our methodology of research programmes shows the rationality of this attitude but it also shows the irrationality of the defence of such 'malformations' once the progressive phase is over.

It should be said here that in the thirties and forties Bohr abandoned his demand for 'new phenomena' and was prepared to 'proceed with the immediate task of co-ordinating the multifarious evidence regarding atomic phenomena, which accumulated from day to day in the exploration of this new field of knowledge'.[1] This indicates that Bohr, by this time, had fallen back on 'saving the phenomena', while Einstein sarcastically insisted that 'every theory is true provided that one suitably associates its symbols with observed quantities'.[2])

But *consistency*—in a strong sense of the term[3]—*must remain an important regulative principle* (over and above the requirement of progressive problemshift); and inconsistencies *must* be seen as problems. The reason is simple. If science aims at truth, it must aim at consistency; if it resigns consistency, it resigns truth. To claim that 'we must be modest in our demands',[4] that we must resign ourselves to—weak or strong—inconsistencies, remains a methodological vice. On the other hand, this does not mean that the discovery of an inconsistency—or of an anomaly—must *immediately* stop the development of a programme: it may be rational to put the inconsistency into some temporary, *ad hoc* quarantine, and carry on with the positive heuristic of the programme. This has been done even in mathematics, as the examples of the early infinitesimal calculus and of naive set theory show.[5]

[1] Bohr [1949], p. 206.

[2] Quoted in Schrödinger [1958], p. 170.

[3] Two propositions are inconsistent if their conjunction has no model, that is, there is no interpretation of their descriptive terms in which the conjunction is true. But in informal discourse we use more formative terms than in formal discourse: some descriptive terms are given a fixed interpretation. In this informal sense two propositions may be (weakly) inconsistent given the standard interpretations of some characteristic terms even if formally, in some unintended interpretation, they may be consistent. For instance, the first theories of electron spin were inconsistent with the special theory of relativity if 'spin' was given its ('strong') standard interpretation and thereby treated as a formative term; but the inconsistency disappears if 'spin' is treated as an uninterpreted descriptive term. The reason why we should not give up standard interpretations too easily is that such emasculation of meanings may emasculate the positive heuristic of the programme. (On the other hand, such meaning shifts may be in some cases progressive: cf. *above*, p. 126.)

For the shifting demarcation between formative and descriptive terms in informal discourse, cf. my [1963–4], 9(*b*), especially p. 335, footnote 1.

[4] Bohr [1922], last paragraph.

[5] Naive falsificationists tend to regard this liberalism as a *crime against reason.* Their main argument runs like this: 'If one were to accept contradictions, then one would have to give up any kind of scientific activity: it would mean a complete breakdown of science. This can be shown by proving that *if two contradictory statements are admitted, any*

(From this point of view, Bohr's 'correspondence principle' played an interesting double role in his programme. On the one hand it functioned as an important heuristic principle which suggested many new scientific hypotheses which, in turn, led to novel facts, especially in the field of the intensity of spectrum lines.[1] On the other hand it functioned also as a defence-mechanism, which 'endeavoured to utilize to the utmost extent the concepts of the classical theories of mechanics and electrodynamics, in spite of the contrast between these theories and the quantum of action',[2] instead of emphasizing the urgency of a unified programme. In this second role it reduced the degree of problematicality of the programme.[3])

Of course, the research programme of quantum theory as a whole was a 'grafted programme' and therefore repugnant to physicists with deeply conservative views like Planck. There are two extreme and equally irrational positions with regard to a grafted programme.

The *conservative position* is to halt the new programme until the basic inconsistency with the old programme is somehow repaired: it is irrational to work on inconsistent foundations. The 'conservatives' will concentrate on eliminating the inconsistency by explaining (approximately) the postulates of the new programme in terms of the old programme: they find it irrational to go on with the new programme without a successful *reduction* of the kind mentioned. Planck himself chose this way. He did not succeed, in spite of the decade of hard work he invested in it.[4] Therefore Laue's remark that his lecture on 14 December 1900, was the 'birthday of the quantum theory' is not quite true: that day was the birthday of Planck's reduction programme. The decision to go *ahead* with temporarily inconsistent foundations was taken by Einstein in 1905, but even he wavered in 1913, when Bohr forged forward again.

statement whatever must be admitted; for from a couple of contradictory statements any statement whatever can be validly inferred . . . A theory which involves a contradiction is therefore entirely useless *as a theory* ' (Popper [1940]). In fairness to Popper, one has to stress that he is here arguing against Hegelian dialectic, in which inconsistency becomes a *virtue*; and he is absolutely right when he points out its dangers. But Popper never analysed patterns of empirical (or non-empirical) progress on inconsistent foundations; indeed, in section 24 of his [1934] he makes consistency and falsifiability mandatory requirements for any scientific theory. I discuss this problem in more detail in my [1970].

[1] Cf. e.g. Kramers [1923].

[2] Bohr [1923].

[3] Born, in his [1954], gives a vivid account of the correspondence principle which strongly supports this double appraisal: 'The art of guessing correct formulae, which deviate from the classical ones, yet contain them as a limiting case . . . was brought to a high degree of perfection.'

[4] For the fascinating story of this long series of frustrating failures, cf. Whittaker, [1953], pp. 103–4. Planck himself gives a dramatic description of these years: 'My futile attempts to fit the elementary quantum of action into the classical theory continued for a number of years, and they cost me a great deal of effort. Many of my colleagues saw in this something bordering on a tragedy . . .' (Planck [1947]).

The *anarchist position* concerning grafted programmes is to extol anarchy in the foundations as a virtue and regard [weak] inconsistency either as some basic property of nature or as an ultimate limitation of human knowledge, as some of Bohr's followers did.

The *rational position* is best characterized by Newton's, who faced a situation which was to a certain extent similar to the one discussed. Cartesian push-mechanics, on which Newton's programme was originally grafted, was (weakly) inconsistent with Newton's theory of gravitation. Newton worked both on his positive heuristic (successfully) *and* on a reductionist programme (unsuccessfully), and disapproved both of Cartesians who, like Huyghens, thought that it was not worth wasting time on an ' unintelligible' programme and of some of his rash disciples who, like Cotes, thought that the inconsistency presented no problem.[1]

The rational position with regard to 'grafted' programmes is then to exploit their heuristic power without resigning oneself to the fundamental chaos on which it is growing. On the whole, this attitude dominated old, pre-1925 quantum theory. In the new, post-1925 quantum theory the 'anarchist' position became dominant and modern quantum physics, in its 'Copenhagen interpretation', became one of the main standard bearers of philosophical obscurantism. In the *new* theory Bohr's notorious 'complementarity principle' enthroned [weak] inconsistency as a basic factual final feature of nature, and merged subjectivist positivism and antilogical dialectic and even ordinary language philosophy into one unholy alliance. After 1925 Bohr and his associates introduced a new and unprecedented lowering of critical standards for scientific theories. This led to a defeat of reason within modern physics and to an anarchist cult of incomprehensible chaos. Einstein protested: 'The Heisenberg–Bohr tranquillizing philosophy—or religion?—is so delicately contrived that, for the time begin, it provides a gentle pillow for the true believer'.[2] On the other hand,

[1] Cf. my [1970]. Of course, a reductionist programme is scientific only if it explains more than it has set out to explain; otherwise the reduction is *not* scientific (cf. Popper [1969]). If the reduction does not produce new empirical content, let alone novel facts, then the reduction represents a degenerating problemshift—it is a mere linguistic exercise. The Cartesian efforts to bolster up their metaphysics in order to be able to interpret Newtonian gravitation in its terms, is an outstanding example for such merely linguistic reduction. Cf. *above*, p. 126, footnote 2.

[2] Einstein [1928]. Among the critics of the Copenhagen 'anarchism' we should mention —besides Einstein—Popper, Landé, Schrödinger, Margenau, Blokhinzev, Bohm, Fényes and Jánossy. For a defence of the Copenhagen interpretation, cf. Heisenberg [1955]; for a hard-hitting recent criticism, cf. Popper [1967]. Feyerabend in his [1968–9], makes use of some inconsistencies and waverings in Bohr's position for a crude apologetic falsification of Bohr's philosophy. Feyerabend misrepresents Popper's, Landé's and Margenau's critical attitude to Bohr, gives insufficient emphasis to Einstein's opposition, and seems to have forgotten completely that in some of his earlier papers he was more Popperian than Popper on this issue.

Einstein's *too* high standards may well have been the reason that prevented him for discovering (or perhaps only from publishing) the Bohr model and wave mechanics.

Einstein and his allies have not won the battle. Physics textbooks are nowadays full of statements like this: 'The two viewpoints, quanta and electromagnetic field strengths, are complementary in the sense of Bohr. This complementarity is one of the great achievements of natural philosophy in which the Copenhagen interpretation of the epistemology of quantum theory has resolved the age-old conflict between the corpuscular and the wave theories of light. From the reflection and rectilinear propagation properties of Hero of Alexandria in the first century A.D., right through to the interference and wave properties of Young and Maxwell in the nineteenth century, this controversy raged. The quantum theory of radiation during the past half century, in a striking Hegelian manner, has *completely* resolved the dichotomy'.[1]

Let us now return to the logic of discovery of *old* quantum theory and, in particular, concentrate on its *positive heuristic*. Bohr's plan was to work out first the theory of the hydrogen atom. His first model was to be based on a fixed proton-nucleus with an electron in a circular orbit; in his second model he wanted to calculate an elliptical orbit in a fixed plane; then he intended to remove the clearly artificial restrictions of the fixed nucleus and fixed plane; after this he thought of taking the possible spin of the electron into account,[2] and then he hoped to extend his programme to the structure of complicated atoms and molecules and to the effect of electromagnetic fields on them, etc., etc. All this was planned right at the start: the idea that atoms are analogous to planetary systems adumbrated a long, difficult but optimistic programme and clearly indicated the policy of research.[3] 'It looked at this time—in the year 1913—as if the authentic

[1] Power [1964], p. 31 (my italics). '*Completely*' is meant here literally. As we read in *Nature* (222, 1969, pp. 1034–5): 'It is absurd to think that any fundamental element of [quantum] theory can be false . . . The arguments that *scientific* results are always temporary, cannot hold. It is the *philosophers*' conceptions of modern physics that are temporary, because they have not yet realized how profoundly the discoveries of quantum physics affect the whole of epistemology. . . . The assertion that ordinary language is the ultimate source of the unambiguousness of physical description is verified most convincingly by the observational conditions in quantum physics.'

[2] This is rational reconstruction. As a matter of fact, Bohr accepted this idea only in his [1926].

[3] Besides this analogy, there was another basic idea in Bohr's positive heuristic: the 'correspondence principle'. This was indicated by him as early as 1913 (cf. the second of his five postulates quoted above on p. 141), but he developed it only later when he used it as a guiding principle in solving some problems of the later, sophisticated models (like the intensities and states of polarization). The peculiarity of this second part of his positive heuristic was that Bohr did not believe its metaphysical version: he thought it was a temporary rule until the replacement of classical electromagnetics (and possibly mechanics).

key to the spectra had at last been found, as if only time and patience would be needed to resolve their riddles completely.'[1]

Bohr's celebrated first paper of 1913 contained the initial step in the research programme. It contained his first model (I shall call it M_1) which already predicted facts hitherto unpredicted by any previous theory: the wavelengths of hydrogen's line emission spectrum. Though some of these wavelengths were known before 1913—the Balmer series (1885) and the Paschen series (1908)—Bohr's theory predicted much more than these two known series. And tests soon corroborated its novel content: one additional Bohr series was discovered by Lyman in 1914, another by Brackett in 1922 and yet another by Pfund in 1924.

Since the Balmer and the Paschen series were known before 1913, some historians present the story as an example of a Baconian 'inductive ascent': (1) the chaos of spectrum lines, (2) an 'empirical law' (Balmer), (3) the theoretical explanation (Bohr). This certainly looks like the three 'floors' of Whewell. But the progress of science would hardly have been delayed had we lacked the laudible trials and errors of the ingenious Swiss school-teacher: the speculative mainline of science, carried forward by the bold speculations of Planck, Rutherford, Einstein and Bohr would have produced Balmer's results deductively, as test-statements of their theories, without Balmer's so-called 'pioneering'. In the rational reconstruction of science there is little reward for the pains of the discoverers of 'naive conjectures'.[2]

As a matter of fact, Bohr's problem was not to explain Balmer's and Paschen's series, but to explain the paradoxical stability of the Rutherford atom. Moreover, Bohr had not even heard of these formulae before he wrote the first version of his paper.[3]

Not all the novel content of Bohr's first model M_1 was corroborated. For instance, Bohr's M_1 claimed to predict all the lines in the hydrogen emission spectrum. But there was experimental evidence for a hydrogen

[1] Davisson [1937]. A similar euphoria was experienced by MacLaurin in 1748 over Newton's programme: Newton's 'philosophy being founded on experiment and demonstration, cannot fail till reason or the nature of things are changed . . . [Newton] left to posterity little more to do, but observe the heavens, and compute after his models' (MacLaurin [1748], p. 8).

[2] I use here 'naive conjecture' as a technical term in the sense of my [1963–4]. For a case study and detailed criticism of the myth of the 'inductive basis' of science (natural or mathematical) cf. *ibid.* section 7, especially pp. 298–307. There I show that Descartes's and Euler's 'naive conjecture' that for all polyhedra $V-E+F=2$ was irrelevant and superfluous for the later development; as further examples one may mention that Boyle's and his successors' labours to establish $pv = RT$ was irrelevant for the later theoretical development (except for developing some experimental techniques), as Kepler's three laws may have been superfluous for the Newtonian theory of gravitation.

For further discussion of this point cf. *below*, p. 175.

[3] Cf. Jammer [1966], pp. 77 ff.

series where according to Bohr's M_1 there should have been none. The anomalous series was the Pickering–Fowler ultraviolet series.

Pickering discovered this series in 1896 in the spectrum of the star ζ Puppis. Fowler, after having discovered its first line also in the sun in 1898, produced the whole series in a discharge tube containing hydrogen and helium. True, it could be argued that the monster-line had nothing to do with the hydrogen—after all, the sun and ζ Puppis contain many gases and the discharge tube also contained helium. Indeed, the line could *not* be produced in a pure hydrogen tube. But Pickering's and Fowler's 'experimental technique', that led to a falsifying hypothesis of Balmer's law, had a plausible, although never severely tested, theoretical background: (*a*) their series had the same convergence number as the Balmer series and therefore was taken to be a hydrogen series and (*b*) Fowler gave a plausible explanation why helium could not possibly be responsible for producing the series.[1]

Bohr was not, however, very impressed by the 'authoritative' experimental physicists. He did not question their 'experimental precision' or the 'reliability of their observations', but questioned their observational theory. Indeed, he proposed an alternative. He first elaborated a new model (M_2) of his research programme: the model of ionized helium, with a double proton orbited by an electron. Now this model predicts an ultraviolet series in the spectrum of ionized helium which coincides with the Pickering–Fowler series. This constituted a rival theory. Then he suggested a 'crucial experiment': he predicted that Fowler's series can be produced, possibly with even stronger lines, in a tube which is filled with a mixture of helium and chlorine. Moreover, Bohr explained to the experimentalists, without even looking at their apparatus, the catalytic role of the hydrogen in Fowler's experiment and of chlorine in the experiment he suggested.[2] Indeed, he was right.[3] Thus the first apparent defeat of the research programme was turned into a resounding victory.

[1] Fowler [1912]. Incidentally his 'observational' theory was provided by 'Rydberg's theoretical investigations' which 'in the absence of strict experimental proof [he] regarded as justifying [his experimental] conclusion' (p. 65). But his theoretician colleague, Professor Nicholson, referred three months later to Fowler's findings as 'laboratory confirmations of Rydberg's theoretical deduction' (Nicholson [1913]). This little story, I think, bears out my pet thesis that most scientists tend to understand little more *about* science than fish about hydrodynamics.

In the Report of the Council to the Ninety-third Annual General Meeting of the Royal Astronomical Society, Fowler's 'observation in laboratory experiments' of new 'hydrogen lines which have so long eluded the efforts of the physicists' is described as 'an advance of great interest' and as 'a triumph of well-directed experimental work'. [2] Bohr [1913*b*].

[3] Evans [1913]. For a similar example of a theoretical physicist teaching a refutation-keen experimentalist what he—the experimentalist—had really observed, cf. *above*, p. 130, footnote 5.

The victory, however, was immediately questioned. Fowler acknowledged that his series was not a hydrogen, but a helium series. But he pointed out that Bohr's monster-adjustment[1] still failed: the wavelengths in the Fowler series differ significantly from the values predicted by Bohr's M_2. Thus the series, although it does not refute M_1, still refutes M_2, and because of the close connection between M_1 and M_2, it undermines M_1![2]

Bohr brushed off Fowler's argument: *of course* he never meant M_2 to be taken too seriously. His values were based on a crude calculation based on the electron orbiting round a fixed nucleus; but *of course* it orbits round the common centre of gravity; *of course*, as is done when treating two-body problems, one has to substitute reduced mass for mass: $m_e' = m_e/[1+(m_e/m_n)]$.[3] This modified model was Bohr's M_3. And Fowler himself had to admit that Bohr was again right.[4]

The apparent refutation of M_2 turned into a victory for M_3; and it was clear that M_2 and M_3 would have been developed within the research programme—perhaps even M_{17} or M_{20}—without *any* stimulus from observation or experiment. It was at this stage that Einstein said of Bohr's theory: 'It is one of the greatest discoveries.'[5]

Bohr's research programme then went on as planned. The next step was to calculate elliptical orbits. This was done by Sommerfeld in 1915, but with the (unexpected) result that the increased number of possible steady orbits did *not* increase the number of possible energy levels, so there seemed to be no possibility of a crucial experiment between the elliptical and circular theory. However, electrons orbit the nucleus with very high velocity so that when they accelerate their mass should change noticeably if Einsteinian mechanics is true. Indeed, calculating such relativistic corrections, Sommerfeld got a new array of energy levels and thus the 'finestructure' of the spectrum.

The switch to this new relativistic model required much more mathematical skill and talent than the development of the first few models. Sommerfeld's achievement was primarily mathematical.[6]

[1] Monster-adjustment: turning a counterexample, in the light of some new theory, into an example. Cf. my [1963–4], pp. 127 ff. But Bohr's 'monster-adjustment' was empirically 'progressive': it predicted a new fact (the appearance of the 4686 line in tubes containing no hydrogen). [2] Fowler [1913*a*].

[3] Bohr [1913*c*]. This monster-adjustment was also 'progressive': Bohr predicted that Fowler's observations must be slightly imprecise and the Rydberg 'constant' must have a fine structure.

[4] Fowler [1913*b*]. But he sceptically noted that Bohr's programme had not yet explained the spectrum lines of *un-ionized*, ordinary helium. However, he soon abandoned his scepticism and joined Bohr's research programme (Fowler [1914]).

[5] Cf. Hevesy [1913]: 'When I told him of the Fowler spectrum, the big eyes of Einstein looked still bigger and he told me: "Then it is one of the greatest discoveries." '

[6] For the vital mathematical aspects of research programmes, cf. *above*, p. 137.

Curiously, the doublets of the hydrogen spectrum had already been discovered in 1891 by Michelson.[1] Moseley pointed out immediately after Bohr's first publication that 'it fails to account for the second weaker line found in each spectrum'.[2] Bohr was not upset: he was convinced that the positive heuristic of his research programme would, *in due course*, explain and even correct Michelson's observations.[3] And so it did. Sommerfeld's theory was, of course, inconsistent with Bohr's first versions; the fine-structure experiments—with the old observations corrected!—provided the crucial evidence in its favour. Many defeats of Bohr's first models were turned by Sommerfeld and his Munich school into victories for Bohr's research programme.

It is interesting that just as Einstein got worried and slowed down in the middle of the spectacular progress of quantum physics by 1913, Bohr got worried and slowed down by 1916; and just as Bohr had, by 1913 taken the initiative from Einstein, Sommerfeld had taken the initiative from Bohr by 1916. The difference between the atmosphere of Bohr's Copenhagen school and Sommerfeld's Munich school was conspicuous: 'In Munich one used more concrete formulations and was therefore more easily understood; one had been successful in the systematization of spectra and in the use of the vector model. In Copenhagen, however, one believed that an adequate language for the new [phenomena] had not yet been found, one was reticent in the face of too definite formulations, one expressed oneself more cautiously and more in general terms, and was therefore much more difficult to understand.'[4]

Our sketch shows how a progressive shift may lend credibility—and *rationale*—to an inconsistent programme. Born, in his obituary of Planck, describes this process forcefully: 'Of course the mere introduction of the quantum of action does not yet mean that a *true* Quantum Theory has been established . . . The difficulties which the introduction of the quantum of action into the well-established classical theory has encountered from the outset have already been indicated. They have gradually increased rather than diminished; and although research in its forward march has in the meantime passed over some of them, the remaining gaps

[1] Michelson [1891–2], especially pp. 287–9. Michelson does not even mention Balmer.

[2] Moseley [1914].

[3] Sommerfeld [1916], p. 68.

[4] Hund [1961]. This is discussed at some length in Feyerabend [1968–9], pp. 83–7 But Feyerabend's paper is heavily biased. The main aim of his paper is to play down Bohr's methodological anarchism and show that Bohr *opposed* the Copenhagen interpretation of the *new* (post-1925) quantum programme. In order to do so, Feyerabend, on the one hand, overemphasizes Bohr's unhappiness about the inconsistency of the *old* (pre-1925) quantum programme and, on the other hand, makes too much of the fact that Sommerfeld cared less for the problematicality of the inconsistent foundations of the *old* programme than Bohr.

in the theory are the more distressing to the conscientious theoretical physicist. In fact, what in Bohr's theory served as the basis of the laws of action consists of certain hypotheses which a generation ago would doubtless have been flatly rejected by every physicist. That within the atom certain quantized orbits (i.e. picked out on the quantum principle) should play a special role could well be granted; somewhat less easy to accept is the further assumption that the electrons moving on these curvilinear orbits, and therefore accelerated, radiate no energy. But that the sharply defined frequency of an emitted light quantum should be different from the frequency of the emitting electron would be regarded by a theoretician who had grown up in the classical school as monstrous and almost inconceivable. But numbers [or, rather, *progressive problemshifts*] decide, and in consequence the tables have been turned. While originally it was a question of fitting in with as little strain as possible a new and strange element into an existing system which was generally regarded as settled, *the intruder, after having won an assured position, now has assumed the offensive*; and it now appears certain that it is about to blow up the old system at some point. The only question now is, at what point and to what extent this will happen.'[1]

One of the most important points one learns from studying research programmes is that relatively few experiments are really important. The heuristic guidance the theoretical physicist receives from tests and 'refutations' is usually so trivial that large-scale testing—or even bothering too much with the data already available—may well be a waste of time. In most cases we need no refutations to tell us that the theory is in urgent need of replacement: the positive heuristic of the programme drives us forward anyway. Also, to give a stern 'refutable interpretation' to a fledgling version of a programme is dangerous methodological cruelty. The first versions may even 'apply' only to non-existing 'ideal' cases; it may take decades of theoretical work to arrive at the first novel facts and still more time to arrive at *interestingly testable* versions of the research programmes, at the stage when refutations are no longer forseeable in the light of the programme itself.

The dialectic of research programmes is then not necessarily an alternating series of speculative conjectures and empirical refutations. The interaction between the development of the programme and the empirical checks may be very varied—which pattern is actually realized depends only on historical accident. Let us mention three typical variants.

(1) Let us imagine that each of the first three consecutive versions, H_1, H_2, H_3 predict some new facts successfully but others unsuccessfully, that is

[1] Born [1948], p. 180; my italics.

each version is both corroborated *and* refuted in turn. Finally H_4 is proposed which predicts some novel facts but stands up to the severest tests. The problemshift is progressive, and also we have a beautiful Popperian alternation of conjectures and refutations.[1] People will admire this as a classical example of theoretical and experimental work going hand in hand.

(2) Another pattern could have been a lone Bohr (possibly without Balmer preceding him), working out H_1, H_2, H_3, H_4 but self-critically withholding publication until H_4. Then H_4 is tested: all the evidence will turn up as corroborations of H_4, the first (and only) published hypothesis. The theoretician—at his desk—is here seen to work far ahead of the experimenter: we have a period of relative autonomy of theoretical progress.

(3) Let us now imagine that *all* the empirical evidence mentioned in these three patterns is already there at the time of the invention of H_1, H_2, H_3, H_4. In this case H_1, H_2, H_3, H_4 will not represent an empirically progressive problemshift and therefore, although all the evidence supports his theories, the scientist has to work on further in order to prove the scientific value of his programme.[2] Such a state of affairs may be brought about either by the fact that an older research programme (which has been challenged by the one leading to H_1, H_2, H_3, H_4) had already produced all these facts—or by the fact that too much government money lay around for collecting data about spectrum lines and hacks stumbled upon all the data. However the latter case is extremely unlikely, for, as Cullen used to say, 'the number of false facts, afloat in the world, infinitely exceeds that of the false theories'[3]; in most such cases the research programme will clash with the available 'facts', the theoretician will look into the 'experimental techniques' of the experimentalist, and having overthrown and replaced his observational theories will correct his facts thereby producing *novel* ones.[4]

[1] In the first three patterns we do not involve complications like successful appeals against the verdict of the experimental scientists.

[2] This shows that if exactly the same theories and the same evidence is rationally reconstructed in different time orders, they may constitute either a progressive or a degenerative shift. Also cf. my [1968*a*], p. 387.

[3] Cf. McCulloch [1825], p. 21. For a strong argument on how extremely unlikely such a pattern is, see *below*, pp. 156–7.

[4] Perhaps it should be mentioned that manic data collection—and 'too much' precision—prevents even the formation of naive 'empirical' hypotheses like Balmer's. Had Balmer known of Michelson's fine-spectra, would he have ever found his formula? Or, had Tycho de Brahe's data been more precise, would Kepler's elliptical law ever have been put forward? The same applies to the naive first version of the general gas law, etc. The Descartes–Euler conjecture on polyhedra might never have been made but for the scarcity of data; cf. my [1963–4], pp. 298 ff.

After this methodological excursion, let us return to Bohr's programme. Not all developments in the programme were foreseen and planned when the positive heuristic was first sketched. When some curious gaps appeared in Sommerfeld's sophisticated models (some predicted lines never did appear), Pauli proposed a deep auxiliary hypothesis (his 'exclusion principle') which accounted not only for the known gaps but reshaped the shell theory of the periodic system of elements and anticipated facts then unknown.

I do not wish to give here an elaborate account of the development of Bohr's programme. But its detailed study from the methodological viewpoint is a veritable goldmine: its marvellously fast progress—on inconsistent foundations!—was breathtaking, the beauty, originality and empirical success of its auxiliary hypotheses, put forward by scientists of brilliance and even genius, was unprecedented in the history of physics.[1] Occasionally the next version of the programme required only a trivial improvement, like the replacement of mass by reduced mass. Occasionally, however, to arrive at the next version required new sophisticated mathematics, like the mathematics of the many-body problem, or new sophisticated physical auxiliary theories. The additional mathematics or physics was either dragged in from some part of extant knowledge (like relativity theory) or invented (like Pauli's exclusion principle). In the latter case we have a 'creative shift' in the positive heuristic.

But even this great programme came to a point where its heuristic power petered out. *Ad hoc* hypotheses multiplied and could not be replaced by content-increasing explanations. For instance, Bohr's theory of molecular (band) spectra predicted the following formula for diatomic molecules:

$$\nu = \frac{h}{8\pi^2 I}[(m+1)^2 - m^2]$$

But the formula was refuted. Bohrians replaced the term m^2 by $m(m+1)$: this fitted the facts but was sadly *ad hoc*.

Then came the problem of some unexplained doublets in alkali spectra. Landé explained them in 1924 by an *ad hoc* 'relativistic splitting rule', Goudsmit and Uhlenbeck in 1925 by electron spin. If Landé's explanation was *ad hoc*, Goudsmit's and Uhlenbeck's was also inconsistent with special relativity theory: surface points on the largish electron had to travel

[1] 'Between the appearance of Bohr's great trilogy in 1913 and the advent of wave mechanics in 1925, a large number of papers appeared developing Bohr's ideas into an impressive theory of atomic phenomena. It was a collective effort and the names of the physicists contributing to it make up an imposing roll-call: Bohr, Born, Klein, Rosseland, Kramers, Pauli, Sommerfeld, Planck, Einstein, Ehrenfest, Epstein, Debye, Schwarzschild, Wilson . . .' (Ter Haar [1967], p. 43).

faster than light, and the electron had even to be bigger than the whole atom.[1] Considerable courage was needed to propose it. (Kronig got the idea earlier but refrained from publishing it because he thought it was inadmissible.[2])

But temerity in proposing wild inconsistencies did not reap any more rewards. The programme lagged behind the discovery of 'facts'. Undigested anomalies swamped the field. With ever more sterile inconsistencies and ever more *ad hoc* hypotheses, the degenerating phase of the research programme had set in: it started—to use one of Popper's favourite phrases —'to lose its empirical character'.[3] Also many problems, like the theory of perturbations, could not even be expected to be solved within it. A rival research programme soon appeared: wave mechanics. Not only did the new programme, even in its first version (de Broglie, 1924), explain Planck's and Bohr's quantum conditions; it also led to an exciting new fact, to the Davisson–Germer experiment. In its later, ever more sophisticated versions it offered solutions to problems which had been completely out of the reach of Bohr's research programme, and explained the *ad hoc* later theories of Bohr's programme by theories satisfying high methodological standards. Wave mechanics soon caught up with, vanquished and replaced Bohr's programme.

De Broglie's paper came at the time when Bohr's programme was degenerating. But this was mere coincidence. One wonders what would have happened if de Broglie had written and published his paper in 1914 instead of 1924.

(*d*) *A new look at crucial experiments: the end of instant rationality.*

It would be wrong to assume that one must stay with a research programme until it has exhausted all its heuristic power, that one must not introduce a rival programme before everybody agrees that the point of degeneration has probably been reached. (Although one can understand the irritation of a physicist when, in the middle of the progressive phase of a research programme, he is confronted by a proliferation of vague metaphysical theories

[1] A footnote in their paper reads: 'It should be observed that [according to our theory] the peripheral velocity of the electron would considerably exceed the velocity of light' (Uhlenbeck and Goudsmit [1925]).

[2] Jammer [1966], pp. 146–8 and 151.

[3] For a vivid description of this degenerating phase of Bohr's programme, cf. Margenau [1950], pp. 311–3.

In the progressive phase of a programme the main heuristic stimulus comes from the positive heuristic: anomalies are largely ignored. In the degenerating phase the heuristic power of the programme peters out. In the absence of a rival programme this situation may be reflected in the psychology of the scientists by an unusual hypersensitivity to anomalies and by a feeling of a Kuhnian 'crisis'.

stimulating no empirical progress.[1]) One must never allow a research programme to become a *Weltanschauung*, or a sort of *scientific rigour*, setting itself up as an arbiter between explanation and non-explanation, as mathematical rigour sets itself up as an arbiter between proof and non-proof. Unfortunately this is the position which Kuhn tends to advocate: indeed, what he calls 'normal science' is nothing but a research programme that has achieved monopoly. But, as a matter of fact, research programmes have achieved complete monopoly only rarely and then only for relatively short periods, in spite of the efforts of some Cartesians, Newtonians and Bohrians. *The history of science has been and should be a history of competing research programmes (or, if you wish, 'paradigms'), but it has not been and must not become a succession of periods of normal science: the sooner competition starts, the better for progress.* 'Theoretical pluralism' is better than 'theoretical monism': on this point Popper and Feyerabend are right and Kuhn is wrong.[2]

The idea of competing scientific research programmes leads us to the problem: *how are research programmes eliminated?* It has transpired from our previous considerations that a degenerating problemshift is no more a sufficient reason to eliminate a research programme than some old-fashioned 'refutation' or a Kuhnian 'crisis'. *Can there be any objective* (as opposed to socio-psychological) *reason to reject a programme, that is, to eliminate its hard core and its programme for constructing protective belts?* Our answer, in outline, is that such an objective reason is provided by a rival research programme which explains the previous success of its rival and supersedes it by a further display of *heuristic power*.[3]

However, the criterion of 'heuristic power' strongly depends on how we construe *'factual novelty'*. Until now we have assumed that it is immediately ascertainable whether a new theory predicts a novel fact or not.[4] But *the novelty of a factual proposition can frequently be seen only after a long period has elapsed.* In order to show this, I shall start with an example.

[1] This is what must have irritated Newton most in the 'sceptical proliferation of theories' by Cartesians. Cf. my [1970].

[2] Nevertheless there is something to be said for at least *some* people sticking to a research programme until it reaches its 'saturation point'; a new programme is then challenged to account for the full success of the old. It is no argument against this that the rival may, when it was first proposed, already have explained all the success of the first programme; the growth of a research programme cannot be predicted—it may stimulate important unforeseeable auxiliary theories of its own. Also, if a version A_n of a research programme P_1 is mathematically equivalent to a version A_m of a rival P_2, one should develop both: their heuristic strength can still be very different.

[3] I use *'heuristic power'* here as a technical term to characterize the power of a research programme to anticipate theoretically novel facts in its growth. I could of course use *'explanatory power'*: cf. *above*, p. 119, footnote 1.

[4] Cf. *above*, p. 116, text to footnote 2, and p. 134, text to footnote 3.

Bohr's theory logically implied Balmer's formula for hydrogen lines as a consequence.[1] Was this a novel fact? One might have been tempted to deny this, since after all, Balmer's formula was well-known. But this is a half-truth. Balmer merely 'observed' B_1: that *hydrogen lines obey the Balmer formula.* Bohr predicted B_2: that *the differences in the energy levels in different orbits of the hydrogen electron obey the Balmer formula.* Now one may say that B_1 already contains all the purely 'observational' content of B_2. But to say this presupposes that there can be a pure 'observational level', untainted by theory, and impervious to theoretical change. In fact, B_1 was accepted only because the optical, chemical and other theories *applied* by Balmer were well corroborated and accepted as interpretative theories; and these theories could always be questioned. It might be argued that we can 'purge' even B_1 of its theoretical presuppositions, and arrive at what Balmer really 'observed', which might be expressed in the more modest assertion, B_0: that *the lines emitted in certain tubes in certain well-specified circumstances (or in the course of a 'controlled experiment'[2]) obey the Balmer formula.* Now some of Popper's arguments show that we can *never* arrive at any hard 'observational' rock-bottom in this way; 'observational' theories can easily be shown to be involved in B_0.[3] On the other hand, given that Bohr's programme after a long progressive development, had shown its heuristic power, its hard core would itself have become well corroborated[4] and therefore qualified as an 'observational' or interpretative theory. But then B_2 will be seen not as a mere theoretical reinterpretation of B_1, but as a *new fact* in its own right.

These considerations lend new emphasis to the hindsight element in our appraisals and lead to a further liberalization of our standards. A new research programme which has just entered the competition may start by explaining 'old facts' in a novel way but may take a very long time before it is seen to produce 'genuinely novel' facts. For instance, the kinetic theory of heat *seemed* to lag behind the results of the phenomenological

[1] Cf. *above*, p. 147.

[2] Cf. *above*, p. 111, footnote 6.

[3] One of Popper's arguments is particularly important: 'There is a widespread belief that the statement "I see that this table here is white", possesses some profound advantage over the statement "This table here is white", from the point of view of epistemology. But from the point of view of evaluating its possible objective tests, the first statement, in speaking about me, does not appear more secure than the second statement, which speaks about the table here' ([1934], section 27). Neurath makes a characteristically blockheaded comment on this passage: 'For us such protocol statements have the advantage of *having more stability*. One may retain the statement: "People in the 16th century saw fiery swords in the sky" while crossing out "There were fiery swords in the sky"' (Neurath [1935], p. 362).

[4] *This remark, incidentally, defines a 'degree of corroboration' for the 'irrefutable' hard cores of research programmes. Newton's theory (in isolation) had no empirical content, yet it was, in this sense, highly corroborated.*

theory for decades before it finally overtook it with the Einstein–Smoluchowski theory of Brownian motion in 1905. After this, what had previously seemed a speculative reinterpretation of old facts (about heat, etc.) turned out to be a discovery of novel facts (about atoms).

All this suggests that we must not discard a budding research programme simply because it has so far failed to overtake a powerful rival. We should not abandon it if, supposing its rival were not there, it would constitute a progressive problemshift.[1] *And we should certainly regard a newly interpreted fact as a new fact, ignoring the insolent priority claims of amateur fact collectors. As long as a budding research programme can be rationally reconstructed as a progressive problemshift, it should be sheltered for a while from a powerful established rival.*[2]

These considerations, on the whole, stress the importance of methodological tolerance, and leave the question of how research programmes are eliminated still unanswered. The reader may even suspect that laying this much stress on fallibility liberalizes or, rather, softens up, our standards to the extent that we will be landed with radical scepticism. Even the celebrated '*crucial experiments*' will then have no force to overthrow a research programme; anything goes.[3]

But this suspicion is unfounded. *Within* a research programme '*minor crucial experiments*' between subsequent versions are quite common. Experiments easily 'decide' between the n-th and $n+1$-th scientific version, since the $n+1$-th is not only inconsistent with the n-th, but also supersedes it. If the $n+1$-th version has more corroborated content in the light of the *same* programme and in the light of the *same* well corroborated observational theories elimination is a relatively routine affair (only relatively, for even here this decision may be subject to appeal). Appeal procedures too are occasionally easy: in many cases the challenged observational theory, far from being well corroborated, is in fact an inarticulate, naive, 'hidden' assumption; it is only the challenge which reveals the existence of this hidden assumption, and brings about its articulation, testing and downfall. Time and again, however, the observational theories are themselves embedded in some research programme and then the appeal procedure

[1] Incidentally, in the methodology of research programmes, the pragmatic meaning of 'rejection' [of a programme] becomes crystal clear: it means *the decision to cease working on it.*

[2] Some might regard—cautiously—this sheltered period of development as '*prescientific*' (or 'theoretical'); and be prepared only when it starts producing 'genuinely novel' facts to recognize its truly *scientific* (or 'empirical') character—but then their recognition will have to be retroactive.

[3] Incidentally, *this conflict between fallibility and criticism can be rightly said to be the main problem—and driving force—of the Popperian research programme in the theory of knowledge.*

leads to a clash between two research programmes: in such cases we may need a '*major crucial experiment*'.

When two research programmes compete, their first 'ideal' models usually deal with different aspects of the domain (for example, the first model of Newton's semi-corpuscular optics described light-refraction, the first model of Huyghens's wave optics light-interference). As the rival research programmes expand, they gradually encroach on each other's territory and the n-th version of the first will be blatantly, dramatically inconsistent with the m-th version of the second.[1] An experiment is repeatedly performed, and as a result, the first is defeated in *this battle*, while the second wins. But *the war* is not over: any research programme is allowed a few such defeats. All its needs for a comeback is to produce an $n+1$-th (or $n+k$-th) content-increasing version and a verification of some of its novel content.

If such a comeback, after sustained effort, is not forthcoming, the war is lost and the original experiment is seen, *with hindsight*, to have been 'crucial'. But especially if the defeated programme is a young, fast-developing programme, and if we decide to give sufficient credit to its 'pre-scientific' successes, allegedly crucial experiments dissolve one after the other in the wake of its forward surge. Even if the defeated programme is an old, established and 'tired' programme, near its 'natural saturation point',[2] it may continue to resist for a long time and hold out with ingenious content-increasing innovations even if these are unrewarded with empirical success. It is very difficult to defeat a research programme supported by talented, imaginative scientists. Alternatively, stubborn defenders of the defeated programme may offer *ad hoc* explanations of the experiments or a shrewd *ad hoc* 'reduction' of the victorious programme to the defeated one. But such efforts we should reject as unscientific.[3]

Our considerations explain why crucial experiments are seen to be crucial only decades later. Kepler's ellipses were generally admitted as crucial evidence for Newton and against Descartes only about one hundred years after Newton's claim. The anomalous behaviour of Mercury's perihelion

[1] An especially interesting case of such competition is *competitive symbiosis*, when a new programme is grafted on to an old one which is inconsistent with it; cf. *above*, p. 142

[2] There is not such thing as a *natural* 'saturation point'; in my [1963–4], especially on pp. 327–8, I was more of a Hegelian, and I thought there was; now I use the expression with an ironical emphasis. There is no predictable or ascertainable limitation on human imagination in inventing new, content-increasing theories or on the 'cunning of reason' (*List der Vernunft*) in rewarding them with some empirical success even if they are false or even if the new theory has less verisimilitude—in Popper's sense—than its predecessor. (Probably all scientific theories ever uttered by men will be false: they still may be rewarded by empirical successes and even have increasing verisimilitude.)

[3] For an example, cf. *above*, p. 126, footnote 2.

was known for decades as one of the many yet unsolved difficulties in Newton's programme; but only the fact that Einstein's theory explained it better transformed a dull anomaly into a brilliant 'refutation' of Newton's research programme.[1] Young claimed that his double-slit experiment of 1802 was a crucial experiment between the corpuscular and the wave programmes of optics; but his claim was only acknowledged much later, after Fresnel developed the wave programme much further 'progressively' and it became clear that the Newtonians could not match its heuristic power. The anomaly, which had been known for decades, received the honorific title of refutation, the experiment the honorific title of 'crucial experiment' only after a long period of uneven development of the two rival programmes. Brownian motion was for nearly a century in the middle of the battlefield before it was *seen* to defeat the phenomenological research programme and turn the war in favour of the atomists. Michelson's 'refutation' of the Balmer series was ignored for a generation until Bohr's triumphant research programme backed it up.

It may be worthwhile to discuss in detail some examples of experiments whose 'crucial' character became evident only retrospectively. First I shall take the celebrated Michelson–Morley experiment of 1887 which allegedly falsified the ether theory and 'led to the theory of relativity', then the Lummer–Pringsheim experiments which allegedly falsified the classical theory of radiation and 'led to the quantum theory'.[2] Finally I shall discuss an experiment which many physicists thought would turn out to decide against the conservation laws but which, in fact, ended up as their most triumphant corroboration.

(*d* 1) *The Michelson–Morley experiment.*

Michelson first devised an experiment in order to test Fresnel's and Stokes's contradictory theories about the influence of the motion of the earth on the ether,[3] during his visit to Helmholtz's Berlin institute in 1881. According to Fresnel's theory, the earth moves through an ether at rest, but the ether within the earth is *partially* carried along with the earth; Fresnel's theory therefore entailed that the velocity of the ether outside the

[1] *Thus an anomaly in a research programme is a phenomenon which we regard as something to be explained in terms of the programme. More generally, we may speak, following Kuhn, about 'puzzles': a 'puzzle' in a programme is a problem which we regard as a challenge to that particular programme. A 'puzzle' can be resolved in three ways: by solving it within the original programme (the anomaly turns into an example); by neutralizing it, i.e. solving it within an independent, different programme (the anomaly disappears); or, finally, by solving it within a rival programme (the anomaly turns into a counterexample).*

[2] Cf. Popper [1934], section 30.

[3] Cf. Fresnel [1818], Stokes [1845] and [1846]. For an excellent brief exposition cf. Lorentz [1895].

earth relative to the earth was positive (i.e. Fresnel's theory implied the existence of an 'ether wind'). According to Stokes's theory, the ether was dragged along by the earth and immediately on the surface of the earth, the velocity of the ether was equal to that of the earth: therefore its relative velocity was zero (i.e. there was *no* ether wind on the surface). Stokes originally thought that the two theories were observationally equivalent: for instance, with suitable auxiliary assumptions both theories explained the aberration of light. But Michelson claimed that his 1881 experiment was a crucial experiment between the two and that it *proved* Stokes's theory.[1] He claimed that the velocity of the earth relative to the ether is far less than Fresnel's theory would have it. Indeed, he concluded that from his experiment' 'the *necessary conclusion* follows that the hypothesis [of a stationary ether] is erroneous. This conclusion *directly contradicts* the explanation of the phenomenon of aberrration which . . . presupposes that the earth moves through the ether, the latter remaining at rest'.[2] As often happens, Michelson the experimenter was then taught a lesson by a theoretician. Lorentz, the leading theoretical physicist of the period, in what Michelson later described as 'a very searching analysis . . . of the entire experiment',[3] showed that Michelson 'misinterpreted' the facts and that what he observed did *not* in fact contradict the hypothesis of the stationary ether. Lorentz showed that Michelson's calculations were wrong; Fresnel's theory predicted only half of the effect Michelson had calculated. Lorentz concluded that Michelson's experiment did *not* refute Fresnel's theory, and that it certainly did not prove Stokes's theory either. Lorentz went on to show that Stokes's theory was inconsistent: that it assumed the ether at the earth's surface to be at rest with regard to the latter *and* required that the relative velocity have a potential; but these two conditions are incompatible. But even if Michelson *had* refuted *one* theory of the stationary ether, the programme is untouched: one can easily devise several other versions of the ether programme, which predict very small values for the ether winds and he, Lorentz, immediately produced one. This theory was testable and Lorentz proudly submitted it to the verdict of experiment.[4] Michelson, jointly with Morley, took up the challenge. The relative velocity of the earth to the ether again seemed to be zero, in conflict with Lorentz's theory. By this time, Michelson had become more cautious in interpreting his data and even thought of the possibility that the solar system as a whole might have moved in the opposite direction to the earth; therefore he decided to repeat the experiment

[1] This transpires, obliquely, from the concluding section of his [1881].
[2] Michelson [1881], p. 128. My italics.
[3] Michelson and Morley [1887], p. 335.
[4] Lorentz [1886]. For the inconsistency of Stokes's theory also cf. his [1892*b*].

'at intervals of three months and thus avoid all uncertainty'.[1] Michelson, in his second paper, does not talk any more about 'necessary conclusions' and 'direct contradictions'. He only thinks that from his experiment 'it appears, from all that precedes, *reasonably certain* that if there be any relative motion between the earth and the luminiferous ether, it must be *small*; quite small enough entirely to refute *Fresnel*'s explanation of aberration'.[2] Thus in this paper Michelson still claims to have refuted Fresnel's theory (and also Lorentz's new theory); but there is not a word about his old 1881 claim that he refuted 'the theory of stationary ether' in general. (Indeed, he believed that in order to do so, he would have to test the ether wind also at high altitudes, 'at the top of an isolated mountain peak, for instance'.[3])

While some ether-theorists—like Kelvin—did not trust Michelson's 'experimental skill',[4] Lorentz pointed out that, in spite of Michelson's naive claim, even his *new* experiment 'furnishes no evidence for the question for which it was undertaken'.[5] One can regard Fresnel's theory perfectly well as an *interpretative* theory, which interprets facts, rather than is refutable by them, and then, Lorentz showed, 'the significance of the Michelson–Morley experiment lies rather in the fact that it can teach us something about *the changes in the dimensions*'[6]: the dimensions of bodies is affected by their movement through the ether. Lorentz elaborated this 'creative shift' within Fresnel's programme with great ingenuity and thereby claimed to have 'removed the contradiction between Fresnel's theory and Michelson's result'.[7] But he admitted that 'since the nature of the molecular forces is entirely unknown to us, it is impossible to test the hypothesis'[8]: *at least for the time being* it could predict no novel facts.[9]

[1] Michelson and Morley [1887], p. 341. But Pearce Williams points out that he never did. (Pearce Williams [1968], p. 34.)
[2] *Ibid.* p. 341. My italics.
[3] Michelson and Morley [1887]. This remark shows that Michelson realized that his 1887 experiment was completely consistent with an ether wind higher up. Max Born, in his [1920], that is, thirty-three years later, asserted that from the 1887 experiment 'we *must* conclude that the ether wind does not exist'. (My italics).
[4] Kelvin said in the 1900 International Congress of Physics that 'the only cloud in the clear sky of the [ether] theory was the null result of the Michelson–Morley experiment' (cf. Miller [1925]) and immediately persuaded Morley and Miller, who were there, to repeat the experiment.
[5] Lorentz [1892*a*].
[6] *Ibid.* My italics.
[7] Lorentz [1895].
[8] Lorentz [1892*b*].
[9] Fitzgerald at the same time, independently of Lorentz, produced a testable version of this 'creative shift' which was quickly refuted by Trouton's, Rayleigh's and Brace's experiments: it was theoretically but not empirically progressive. Cf. Whittaker [1947], p. 53 and Whittaker [1953], pp. 28–30.

There is a widespread view that Fitzgerald's theory was *ad hoc*. What contemporary physicists meant was that the theory was *ad hoc*$_3$ (cf. *above*, p. 125, footnote 1): that there was '*no independent* [*positive*] *evidence*' for it. (Cf. e.g. Larmor [1904], p. 624.) Later,

In the meanwhile, in 1897, Michelson carried out his long planned experiment to measure the velocity of ether wind on mountain tops. He found none. Since he had thought earlier that he had proved Stokes's theory which predicted an ether wind higher up, he was dumbfounded. If Stokes's theory was still correct, the gradient of the velocity of the ether had to be very small. Michelson had to conclude that 'the earth's influence upon the ether extended to distances of the order of the earth's diameter'.[1] He thought that this was an 'improbable' result, and decided that in 1887 he had drawn the wrong conclusion from his experiment: it was Stokes's theory which had to be rejected and Fresnel's which had to be accepted; and he decided that he would accept *any* reasonable auxiliary hypothesis to have it saved, including Lorentz's 1892 theory.[2] He *now* seemed to prefer the Fitzgerald–Lorentz contraction and by 1904 his colleagues at Case were trying to find out whether this contraction varies with different materials.[3]

While most physicists tried to interpret Michelson's experiments within the framework of the ether programme, Einstein, unaware of Michelson, Fitzgerald and Lorentz, but stimulated primarily by Mach's criticism of Newtonian mechanics, arrived at a new, progressive research programme.[4] This new programme not only 'predicted' and explained the outcome of the Michelson–Morley experiment but also predicted a huge array of previously undreamt-of facts, which obtained dramatic corroborations. It was *only then*, twenty-five years later, that the Michelson–Morley experiment came to be seen as 'the greatest negative experiment in the history of science'.[5] But this could not be seen instantly. Even if the experiment was negative, it was not clear, negative exactly to *what*? Moreover, Michelson in 1881 thought that it was also *positive*: he held that he had *refuted* Fresnel's

under Popper's influence the term '*ad hoc*' was primarily used in the sense of *ad hoc*$_1$, that there was *no independent test* possible for it. But, as the refuting experiments show, it is a mistake to claim, as Popper does, that Fitzgerald's theory was *ad hoc*$_1$ (cf. Popper [1934], section 20). This shows again how important it is to separate *ad hoc*$_1$, and *ad hoc*$_2$.

When Grünbaum, in his [1959*a*], pointed out Popper's mistake, Popper admitted it but replied that Fitzgerald's theory was certainly *more ad hoc* than Einstein's (Popper [1959*b*]), and that this provides yet another '. . . excellent example of "degrees of *ad-hocness*" and of one of the main theses of [his] book—that *degrees* of *ad-hocness* are related (inversely) to degrees of testability and significance'. But the difference is *not* simply a matter of degrees of a unique *ad-hocness* which can be measured by testability. Also cf. *below*, p. 175.

[1] Michelson [1897], p. 478.

[2] Lorentz, indeed, immediately commented: 'While [Michelson] considers so far-reaching an influence of the earth improbable, I should, on the contrary, *expect* it' (Lorentz [1897]; my italics).

[3] Morley and Miller [1904].

[4] There has been a considerable controversy about the historico-heuristic background of Einstein's theory, in the light of which this statement may turn out to be false.

[5] Bernal [1965], p. 530. For Kelvin, in 1905, it was only a 'cloud in the clear sky': cf. *above*, p. 161, footnote 4.

but had *verified* Stokes's theory. Michelson himself and then Fitzgerald and Lorentz explained the result also *positively* within the ether programme.[1] As it is with all experimental results, its negativity for the old programme was established *only later*, by the slow accumulation of *ad hoc* attempts to account for it within the degenerating old programme and by the gradual establishment of a new *progressive* victorious programme in which it has become a positive instance. But the possibility of the rehabilitation of some part of the 'degenerating' old programme could never be rationally excluded.

Only an extremely difficult and—indefinitely—long process can establish a research programme as superseding its rival; and it is unwise to use the term 'crucial experiment' too rashly. Even when a research programme is seen to be swept away by its predecessor, it is not swept away by some 'crucial' experiment; and even if some such crucial experiment is later called in doubt, the new research programme cannot be stopped without a powerful progressive upsurge of the old programme.[2] The negativity—and importance—of the Michelson–Morley experiment lies primarily in the progressive shift in the *new* research programme to which it came to lend powerful support, and its 'greatness' is only a reflection of the greatness of the two *programmes* involved.

It would be interesting to give a detailed analysis of the rival shifts involved in the waning fortunes of the ether theory. But under the influence of naive falsificationism the most interesting degenerating phase in the ether theory after Michelson's 'crucial experiment' is simply ignored by most Einsteinians. They believe that the Michelson–Morley experiment single-handedly defeated the ether theory, the tenacity of which was only due to obscurantist conservatism. On the other hand, this post-Michelson period of the ether theory is not scrutinized *critically* by the anti-Einsteinians, who believe that the ether theory suffered no setback whatsoever: what is good in Einstein's theory was essentially in Lorentz's ether theory and Einstein's victory is only due to positivist fashion. But, in fact, Michelson's long series of experiments from 1881 to 1935, conducted in order to test subsequent versions of the ether programme provides a fascinating

[1] Indeed, Chwolson's excellent physics textbook said in 1902 that the probability of the ether hypothesis borders on certainty. (Cf. Einstein [1909], p. 817.)

[2] Polanyi tells us with *gusto* how, in 1925, in his presidential address to the American Physical Society, Miller announced that Michelson's and Morley's reports notwithstanding, he had 'overwhelming evidence' for an ether-drift; yet the audience remained committed to Einstein's theory. Polanyi draws the conclusion that no ' "objectivist" framework' can account for the scientist's acceptance or rejection of theories (Polanyi [1958], pp. 12–14). But my reconstruction makes the tenacity of the Einsteinian research programme in the face of alleged contrary evidence a completely *rational* phenomenon and thereby undermines Polanyi's 'post-critical'-mystical message.

example of a degenerating problemshift.[1] (But problemshifts may get out of degenerating troughs. It is well known that Lorentz's ether theory can easily be strengthened in such a way that it becomes, in an interesting sense, equivalent with Einstein's no-ether theory.[2] The ether may, in the context of a major 'creative shift', still return.[3])

The fact that we heed hindsight to evaluate experiments explains why, between 1881 and 1886, Michelson's experiment was not even mentioned in the literature. Indeed, when a French physicist, Potier, pointed out to Michelson his 1881 mistake, Michelson decided not to publish a correction note. He explains the reason for this decision in a letter to Rayleigh in March 1887: 'I have repeatedly tried to interest my scientific friends in this experiment without avail, and the reason for my never publishing the correction (I am ashamed to confess it) was that I was discouraged at the slight attention the work received, and did not think it worthwhile.'[4] This letter, incidentally, was a reply to a letter from Rayleigh which drew Michelson's attention to Lorentz's paper. This letter triggered off the 1887 experiment. But even after 1887, and even after 1905, the Michelson–Morley experiment was not yet generally regarded as disproving the existence of the ether, and with good reason. This may explain why Michelson was awarded his Nobel Prize (in 1907), not for 'refuting the ether theory', but 'for his optical precision *instruments* and the spectro-scopic and methodological investigations carried out with their aid'[5]; and why the Michelson–Morley experiment was not even mentioned in the presentation speeches. Michelson, in his *Nobel Lecture*, did not mention it; and he kept quiet

[1] *One typical sign of the degeneration of a programme which is not discussed in this paper is the proliferation of contradictory 'facts'. Using a false theory as an interpretative theory, one may get—without committing any 'experimental mistake'—contradictory factual propositions, inconsistent experimental results.* Michelson, who stuck to the ether to the bitter end, was primarily frustrated by the inconsistency of the 'facts' he arrived at by his ultra-precise measurements. His 1887 experiment 'showed' that there was no ether wind on the earth's surface. But aberration 'showed' that there was. Moreover, his own 1925 experiment (either never mentioned or, as in Jaffé's [1960], misrepresented) also 'proved' that there was one (cf. Michelson and Gale [1925] and, for a sharp criticism, Runge [1925]).

[2] Cf. e.g. Ehrenfest [1913], pp. 17–18, quoted and discussed by Dorling in his [1968]. But one should not forget that *two specific theories, while being mathematically (and observationally) equivalent, may still be embedded into different rival research programmes, and the power of the positive heuristic of these programmes may well be different.* This point has been overlooked by proposers of such equivalence proofs (a good example is the equivalence proof between Schrödinger's and Heisenberg's approach to quantum physics). Also cf. *above*, p. 155, footnote 2.

[3] Cf. e.g. Dirac [1951]: 'If one reexamines the question in the light of present-day knowledge, one finds that the aether is no longer ruled out by relativity, and good reasons can now be advanced for postulating an aether.' Also cf. the concluding paragraph of Rabi [1961] and Prokhovnik [1967].

[4] Shankland [1964], p. 29.

[5] My italics.

about the fact that although he might have originally devised his instruments to measure precisely the velocity of light, he was compelled to improve them for testing some specific ether theories and that the 'precision' of his 1887 experiment was largely motivated by Lorentz's theoretical criticism: a fact which standard contemporary literature never mentions.[1]

Finally, one tends to forget that even if the Michelson–Morley experiment had shown an 'ether wind', Einstein's programme might have been victorious nonetheless. When Miller, an ardent champion of the classical ether programme, published his sensational claim that the Michelson–Morley experiment was sloppily conducted and in fact there *was* an ether wind, the news correspondent of *Science* crowed that 'Professor Miller's results knock out the relativity theory radically'.[2] In Einstein's view, however, even if Miller had reported the true state of affairs '[only] the *present form* of relativity theory' would have to be abandoned.[3] In fact, Synge pointed out that Miller's results, even if taken at their face value, do not conflict with Einstein's theory: only Miller's explanation of them does. One can easily replace the extant auxiliary theory of rigid bodies by a new, Gardner–Synge theory, and then Miller's results are fully digested within Einstein's programme.[4]

(*d* 2) *The Lummer–Pringsheim experiments.*

Let us discuss another alleged crucial experiment. Planck claimed that Lummer's and Pringsheim's experiments, which *'refuted'* Wien's and Rayleigh's and Jeans's laws of radiation at the turn of the century, *'led to'*—or 'even brought about'—the quantum theory.[5] But again the role of these experiments is much more complicated and is very much in line with our approach. It is not simply that Lummer's and Pringsheim's experiments put an end to the classical approach but were neatly explained by quantum physics. On the one hand, some early versions of quantum theory by Einstein *entail* Wien's law and therefore were no less refuted by

[1] Einstein himself tended to believe that Michelson devised his interferometer in order to test Fresnel's theory. (Cf. Einstein [1931].) Incidentally, Michelson's early experiments on spectrum lines—like his [1881–2]—were also relevant to the ether theories of his day. Michelson over-emphasized his success in 'precise measurements' only when he was frustrated by his lack of success in evaluating their relevance for theories. Einstein, who disliked precision for its own sake, asked him why he devoted so much energy to it. Michelson's answer was 'because he found it fun.' (Cf. Einstein [1931].)

[2] *Science* [1925].

[3] Einstein [1927]. My italics.

[4] Synge [1952–4].

[5] Planck [1929]. Popper, in his [1934], section 30, Gamow in his [1966] (p. 37), take over this locution. Of course, observation statements do not 'lead' to some uniquely determined theory.

Lummer's and Pringsheim's experiments than the classical theory.[1] On the other hand, several classical explanations of the Planck formula were offered. For instance, at the 1913 meeting of the British Association for the Advancement of Science, there was a special meeting on radiation, attended among others by Jeans, Rayleigh, J. J. Thomson, Larmor, Rutherford, Bragg, Poynting, Lorentz, Pringsheim and Bohr. Pringsheim and Rayleigh were studiedly neutral about quantum theoretical speculations, but Professor Love 'represented the older views, and maintained the possibility of explaining facts about radiation without adopting the theory of quanta. He criticized the application of the equi-partition of energy theory, on which part of the quantum theory rests. The evidence for the quantum theory of most weight is the agreement with experiment of Planck's formula for the emissivity of a black body. From the mathematical point of view, there may be many more formulae which would agree equally well with the experiments. A formula due to A. Korn was dealt with, which gave results over a wide range, showing just about as good agreement with experiment as the Planck formula. In further contention that *the resources of ordinary theory are not exhausted*, he pointed out that it may be possible to extend the calculation for the emissivity of a thin plate due to Lorentz to other cases. For this calculation no simple analytical expression represents the results over the whole range of wavelengths, and it may well be that in the general case no simple formula exists which is applicable to all wavelengths. Planck's formula may, in fact, be nothing more than an empirical formula.'[2] One example of classical explanations was due to Callendar: 'The disagreement with experiment of Wien's well-known formula for the partition of energy in full radiation, is readily explained if we assume that it represents only the intrinsic energy. The corresponding value of the pressure is very easily deduced by reference to Carnot's principle, as Lord Rayleigh has indicated. The formula which I have proposed (*Phil. Mag.*, October 1913) is simply the sum of the pressure and energy-density thus obtained, and gives very satisfactory agreement with experiment, both for radiation and specific heat. I prefer it to Planck's formula (among other reasons) on the ground that the latter cannot be reconciled with the classical thermodynamics, and involves the conception of a *quantum*, or indivisible unit of action, which is unthinkable. The corresponding physical magnitude on my theory, which I have elsewhere called a molecule of caloric, is not necessarily indivisible, but bears a very simple relation to the intrinsic energy of an atom, which is all that is

[1] Cf. Ter Haar [1967], p. 18. A budding research programme usually starts by explaining already refuted 'empirical laws'—and this, in the light of my approach, may be *rationally* regarded as a success.

[2] *Nature* [1913–14], p. 306; my italics.

required to explain the facts that radiation may in special cases be emitted in atomic units which are multiples of a particular magnitude.'[1]

These quotations may have been tediously long but at least they show again convincingly the absence of instant crucial experiments. Lummer's and Pringsheim's refutations did not eliminate the classical approach to the radiation problem. The situation can be better described by pointing out that Planck's original '*ad hoc*' formula[2]—which fitted (and corrected) Lummer's and Pringsheim's data—could be explained *progressively* within the new quantum theoretical programme,[3] while neither his '*ad hoc*' formula, nor its 'semi-empirical' rivals could be explained within the classical programme except at the price of a degenerating problemshift. The 'progressive' development, incidentally, hinged on a 'creative shift': the replacement (by Einstein) of the Boltzman–Maxwell by the Bose–Einstein statistics.[4] The progressiveness of the new development was abundantly clear: in Planck's version it predicted correctly the value of the Boltzman–Planck constant and in Einstein's version it predicted a stunning series of further novel facts.[5] But before the invention of the new—but sadly *ad hoc*—auxiliary hypotheses in the old programme, before the unfolding of the new programme, and before the discovery of the new facts indicating a progressive problemshift in the latter, the objective relevance of the Lummer–Pringsheim experiments was very limited.

[1] Callendar [1914].

[2] I am referring to Planck's formula as given in his [1900*a*] in which he admitted that after having tried for a long time to prove that 'Wien's law must be *necessarily* true', the 'law' was refuted. So he switched from proving lofty eternal laws to 'constructing completely arbitrary expressions'. But of course any physical theory turns out to be 'completely arbitrary' by justificationist standards. In fact, Planck's arbitrary formula contradicted—and victoriously corrected—contemporary empirical evidence. (Planck told this part of the story in his scientific autobiography.) Of course, in an important sense, Planck's *original* radiation formula was 'arbitrary', 'formal', '*ad hoc*': it was a rather isolated formula which was not part of a research programme. (Cf. *below*, p. 175, footnote 3.) As he himself put it: 'Even if the absolutely precise validity of the radiation formula is taken for granted, so long as it had merely the standing of a law disclosed by a lucky intuition, it could not be expected to possess more than a formal significance. For this reason, on the very day when I formulated this law, I began to devote myself to the task of investing it with a true physical meaning' ([1947], p. 41). But the primary importance of 'investing the formula with a physical meaning'—not necessarily '*true* physical meaning'—is that such interpretation frequently leads to a suggestive research programme and *growth*.

[3] First by Planck himself, in his [1900*b*] which 'founded' the research programme of quantum theory.

[4] This had already been done by Planck, but only inadvertently, as it were by mistake. Cf. Ter Haar [1967], p. 18. Indeed, one role of Pringsheim's and Lummer's results was to stimulate the critical analysis of the informal deductions in the quantum theory of radiation, deductions which were loaded with vital 'hidden lemmas' articulated only in the later development. A most important step in this 'articulating process' was Ehrenfest's [1911].

[5] Cf. e.g. Joffé's 1910 list (Joffé [1911], p. 547).

(*d* 3) *Beta-decay versus conservation laws.*

Finally, I shall tell a story of an experiment which very nearly, but not quite, became 'the greatest negative experiment in the history of science'. The story again illustrates the supreme difficulties of deciding exactly *what* one learns from experience, what it 'proves' and what it 'disproves'. The piece of experience under scrutiny will be Chadwick's 'observation' of beta decay in 1914. The story shows how an experiment may first be regarded as presenting a routine puzzle within a research programme, then nearly promoted to the rank of 'crucial experiment', and then again downgraded to presenting a (*new*) routine puzzle, all this depending on the *whole* changing theoretical and empirical landscape. Most conventional accounts are confused by these changes and prefer to falsify history.[1]

When Chadwick discovered the continuous spectrum of radioactive beta-emission in 1914, nobody thought that this curious phenomenon had anything to do with conservation laws. Two ingenious rival explanations were offered in 1922, both within the framework of the atomic physics of the day, one by L. Meitner, the other by C. D. Ellis. According to Miss Meitner, the electrons were partly primary electrons from the nucleus, partly secondary electrons from the electron shell. According to Mr Ellis, they were all primary electrons. Both theories contained sophisticated auxiliary hypotheses, but both predicted novel facts. The predicted facts contradicted each other and the experimental testimony supported Ellis against Meitner.[2] Miss Meitner appealed; the experimental 'appeal court' refused to support her, but ruled that one crucial auxiliary hypothesis in Ellis's theory had to be rejected.[3] The result of the contest was a draw.

Still nobody would have thought that Chadwick's experiment defied the law of conservation of energy, had not Bohr and Kramers arrived exactly at the time of the Ellis–Meitner controversy at the idea that a consistent theory could be developed only if they renounced the principle of conservation of energy in single processes. One of the main features of the fascinating Bohr–Kramers–Slater theory in 1924 was that the classical laws of conservation of energy and momentum were replaced by statistical ones.[4] This theory (or, rather, 'programme') was immediately 'refuted'

[1] A notable partial exception is Pauli's account (Pauli [1958]). In what follows I am trying both to correct Pauli's story and to show that its rationality can be easily seen in the light of our approach.

[2] Ellis and Wooster [1927].

[3] Meitner and Orthmann [1930].

[4] Slater co-operated only reluctantly in sacrificing the conservation principle. He wrote to van der Waerden in 1964: 'As you suspected, the idea of statistical conservation of energy and momentum was put into the theory by Bohr and Kramers, quite against my better judgment.' Van der Waerden does his amusing best to exonerate Slater from the terrible crime of being responsible for a false theory (van der Waerden [1967], p. 13).

and none of its consequences corroborated; indeed, it was never sufficiently developed to explain beta-decay. But in spite of the immediate abandonment of this programme (not simply because of its 'refutations' by the Compton–Simon and Bothe-Geiger experiments but because of the emergence of a powerful rival: the Heisenberg–Schrödinger programme[1]), Bohr remained convinced that the non-statistical conservation laws would finally have to be abandoned and that the beta-decay anomaly would never be explained until these laws were replaced; at which time beta-decay would be seen as a crucial experiment against the conservation laws. Gamow tells us how Bohr tried to use the idea of non-conservation of energy in beta-decay for an ingenious explanation of the seemingly eternal production of energy in stars.[2] Only Pauli, in his Mephistophelian urge to defy the Lord, remained conservative[3] and devised, in 1930, his neutrino theory in order to explain beta decay and in order to save the principle of conservation of energy. He communicated his idea in a jocular letter to a conference in Tübingen—he himself preferred to stay in Zürich to attend a ball.[4] He first mentioned it in a public lecture in 1931 in Pasadena, but he did not allow the lecture to be published because he felt 'unsure' about it. Bohr, at that time (in 1932), still thought that—at least in nuclear physics—one may have 'to renounce the very idea of energy balance'.[5] Pauli finally decided to publish his talk on the neutrino which he delivered to the 1933 Solvay conference, in spite of the fact that 'the reception at the Congress, except for two young physicists, was sceptical'.[6] But Pauli's theory had some methodological merits. It saved not only the principle of conservation of energy but also the principle of conservation of spin and statistics: it explained not only the beta-decay spectrum but, at the same time, the 'nitrogen anomaly'.[7] By Whewellian standards this 'consilience of inductions' should have been sufficient to establish the

[1] Popper is wrong to suggest that these 'refutations' were sufficient to bring about the downfall of this theory. (Popper [1963], p. 242.)

[2] Gamow [1966], pp. 72–4. Bohr never published this theory (it was untestable as it stood) but 'it looked'—writes Gamow—'as if he would not be greatly surprised if it were true'. Gamow does not date this unpublished theory but it seems that Bohr entertained it in 1928–9 when Gamow was working in Copenhagen.

[3] Cf. the amusing play 'Faust' produced in Bohr's institute in 1932; published by Gamow as an appendix to his [1966].

[4] Cf. Pauli [1958], p. 160.

[5] Bohr [1932]. Ehrenfest too sided firmly with Bohr against the neutrino. Chadwick's discovery of the neutron in 1932 only slightly shook their opposition: they still dreaded the idea of a particle which has neither charge nor, possibly, even (rest) mass, but only 'disembodied' spin.

[6] Wu [1966].

[7] For a fascinating discussion of the open problems presented by the beta-decay and by the nitrogen anomaly, cf. Bohr's Faraday Lecture in 1930, read before, but published after, Pauli's solution (Bohr [1930], especially pp. 380–3).

respectability of Pauli's theory. But on our criteria, the successful prediction of some *novel* fact was needed. This too was provided by Pauli's theory. For Pauli's theory had an interesting observable consequence: if it was right, the β-spectra had to have a clear upper bound. This question was *at the time* undecided, but Ellis and Mott became interested[1] and soon, Ellis's student, Henderson, showed that the experiments supported Pauli's programme.[2] Bohr was not impressed. He knew that if a major programme based on *statistical* conservation of energy ever got going, the growing belt of auxiliary hypotheses would take proper care of the most negative-looking evidence.

Indeed, in these years most leading physicists thought that in nuclear physics the laws of conservation of energy and momentum break down.[3] The reason was stated clearly by Lise Meitner who admitted defeat only in 1933: 'All the attempts to uphold the validity of the law of conservation of energy also for *single* processes demanded a second process [in the beta-decay]. But no such process was found . . .'[4]: that is, the conservation programme for the nucleus showed an empirically degenerating problemshift. There were several ingenious attempts to account for the continuous beta-emission spectrum without assuming a 'thief particle'.[5] These attempts were discussed with great interest,[6] but they were abandoned because they failed to establish a progressive shift.

At this point, Fermi entered on the scene. In 1933–4 he reinterpreted the beta-emission problem in the framework of the research programme of the new quantum theory. Thus he initiated a small new research programme of the neutrino (which later grew into the programme of weak interactions). He calculated some first crude models.[7] Although his theory did not yet predict any new fact, he made it clear that this was only a matter of some further work.

Two years passed and Fermi's promise was still not fulfilled. But the new programme of quantum physics developed fast, at least as far as the non-nuclear phenomena were concerned. Bohr became convinced that some of the basic original ideas of the Bohr–Kramers–Slater programme were now firmly embedded in the new quantum programme and that the

[1] Ellis and Mott [1933].
[2] Henderson [1934].
[3] Mott [1933], p. 823. Heisenberg, in his celebrated [1932], in which he introduced the proton-neutron model of the nucleus, pointed out that 'because of the breakdown of the conservation of energy in the beta-decay one cannot give a unique definition of the binding energy of the electron within the neutron' (p. 164).
[4] Meitner [1933], p. 132.
[5] E.g. Thomson [1929] and Kudar [1929–30].
[6] For a most interesting discussion cf. Rutherford, Chadwick and Ellis [1930], pp. 335–6.
[7] Fermi [1933] and [1934].

new programme solved the intrinsic theoretical problems of the old quantum programme without touching the conservation laws. Therefore Bohr followed Fermi's work with sympathy, and in 1936, in an unusual sequence of events, gave it, by our standards prematurely, public support.

In 1936 Shankland devised a new test of rival theories of photon scattering. His results seemed to support the discarded Bohr–Kramers–Slater theory and undermine the reliability of experiments which, more than a decade earlier, refuted it.[1] Shankland's paper created a sensation. Those physicists who abhorred the new trend were quick to hail Shankland's experiment. Dirac, for instance, immediately welcomed back the 'refuted' Bohr–Kramers–Slater programme, wrote a very sharp article against the 'so-called quantum electrodynamics' and demanded 'a profound alteration in current theoretical ideas, involving a departure from the conservation laws [in order] to get a satisfactory relativistic quantum mechanics'.[2] In the article Dirac suggested again that beta-decay may well turn out to be a piece of crucial evidence against the conservation laws and made fun of the 'new unobservable particle, the neutrino, specially postulated by some investigators in an attempt formally to preserve conservation of energy by assuming the unobservable particle to carry off the balance'.[3] Immediately afterwards Peierls joined the discussion. Peierls suggested that Shankland's experiment may turn out to refute even the statistical conservation of energy. He added: 'That, too, seems satisfactory, once detailed conservation has been abandoned.'[4]

In Bohr's Copenhagen institute, Shankland's experiments were immediately repeated and discarded. Jacobsen, a colleague of Bohr reported this in a letter to *Nature*. Jacobsen's results were accompanied by a letter from Bohr himself, who firmly came out against the rebels, and in defence of Heisenberg's new quantum programme. In particular, he came out in defence of the neutrino against Dirac: 'It may be remarked that the grounds for serious doubts as regards the strict validity of the conservation laws in the problem of the emission of β-rays from atomic nuclei are now largely removed by the suggestive agreement between the rapidly increasing experimental evidence regarding β-ray phenomena and the consequences of the neutrino hypotheses of Pauli so remarkably developed in Fermi's theory.'[5]

Fermi's theory, in its first versions, had no striking empirical success. Indeed, even the available data, especially in the case of *RaE*, on which beta emission research then centred, sharply contradicted Fermi's 1933–4 theory. He wanted to deal with these in the second part of his paper which,

[1] Shankland [1936].
[2] Dirac [1936].
[3] Dirac [1936].
[4] Peierls [1936].
[5] Bohr [1936].

however, was never published. Even if one construes Fermi's 1933–4 theory as a first version of a flexible programme, by 1936 one could not possibly detect any serious sign of a progressive shift.[1] But Bohr wanted to put his *authority* behind Fermi's daring application of Heisenberg's new big programme to the nucleus; and since Shankland's experiment and Dirac's and Peierls's attack brought the beta-decay into the focus of the criticism of the new big programme, he over-praised Fermi's neutrino programme which promised to fill in a sensitive gap. No doubt, the later development spared Bohr from a dramatic humilation: the programmes based on conservation principles progressed, while no progress was made in the rival camp.[2]

The moral of this story is again that the status of an experiment as 'crucial' depends on the status of the theoretical competition in which it is embedded. As the fortunes of the competing camps wax or wane, the interpretation and appraisal of the experiment may change.

Our scientific folklore however is impregnated with theories of instant rationality. The story which I described is falsified in most accounts and reconstructed in terms of some wrong theory of rationality. Even the very best popular expositions teem with such falsifications. Let me mention two examples.

In one paper we learn this about beta-decay: 'When this situation was faced for the first time, the alternatives seemed grim. Physicists *either* had to accept a breakdown of the law of energy conservation, *or* they had to suppose the existence of a new and unseen particle. Such a particle, emitted along with the proton and the electron in the disintegration of the neutron, could save the central pillar of physics by carrying off the missing energy. This was in the early 1930s, when the introduction of a new particle was

[1] Several physicists between 1933 and 1936 offered alternatives or proposed *ad hoc* changes of Fermi's theory; cf. e.g. Becke and Sitte [1933], Bethe and Peierls [1934], Konopinski and Uhlenbeck [1934]. Wu and Moszkowski write in 1966 that 'the Fermi theory [i.e. programme] of β-decay is *now* known to predict with remarkable accuracy both the relation between the rate of β-decay and the energy of disintegration, and also the shape of β-spectra'. But they stress that 'at the very beginning the Fermi theory unfortunately met an unfair test. Until the time when artificial radioactive nuclei could be copiously produced, RaE was the only candidate that beautifully fulfilled many experimental requirements as a β source for the investigation of its spectrum shape. How could we have known then that the β spectrum of RaE would turn out to be only a very special case, one whose spectrum has, in fact, been understood only very recently. Its peculiar energy dependence defied what was expected of the simple Fermi theory of β decay and greatly slackened the pace of the theory's [i.e. programme's] initial progress' (Wu and Moszkowski [1966], p. 6).

[2] It is very doubtful whether Fermi's neutrino programme was progressive or degenerating even between 1936 and 1950; and after 1950 the verdict is still not crystal clear. But this I shall try to discuss in some other occasion. (Incidentally, Schrödinger stood up for the statistical interpretation of the conservation principles in spite of his crucial role in the development of new quantum physics; cf. his [1958].)

not the casual matter it is today. Nevertheless, *after only the briefest vacillation*, physicists chose the second alternative.'[1] Of course, even the *discussed* alternatives were many more than two and the 'vacillation' was certainly not 'the briefest'.

In a well-known textbook of philosophy of science we learn that (1) 'the law (or principle) of the conservation of energy was seriously challenged by experiments on beta-ray decay whose outcome could not be denied'; that (2) 'nevertheless, the law was not abandoned, and the existence of a new kind of entity (called a "neutrino") was assumed in order to bring the law into concordance with experimental data'; and that (3) 'the rationale for this assumption is that the rejection of the conservation law would deprive a large part of our physical knowledge of its systematic coherence'.[2] But all the three points are wrong. (1) is wrong because no law can be 'seriously challenged' by experiments only; (2) is wrong because new *scientific* hypotheses are assumed not simply in order to patch up gaps between data and theory but in order to predict novel facts; and (3) is wrong because at the time it seemed that *only* the rejection of the conservation law would secure the 'systematic coherence' of our physical knowledge.

(d 4) Conclusion. The requirement of continuous growth.

There are no such things as crucial experiments, at least not if these are meant to be experiments which can *instantly* overthrow a research programme. In fact, when one research programme suffers defeat and is superseded by another one, we may—*with long hindsight*—call an experiment crucial if it turns out to have provided a spectacular corroborating instance for the victorious programme and a failure for the defeated one (in the sense that it was never 'explained progressively'—or, briefly, 'explained'[3]—within the defeated programme). But scientists, of course, do not always judge heuristic situations correctly. A rash scientist may *claim* that his experiment defeated a programme, and parts of the scientific community may even, rashly, accept his claim. But if a scientist in the 'defeated' camp puts forward a few years later a scientific explanation of the allegedly 'crucial experiment' within (or consistent with) the allegedly defeated programme, *the honorific title may be withdrawn and the 'crucial experiment' may turn from a defeat into a new victory for the programme.*

Examples abound. There were many experiments in the eighteenth century which were, as a matter of historico-sociological fact, widely accepted as 'crucial' evidence against Galileo's law of free fall, and Newton's theory of gravitation. In the nineteenth century there were several 'crucial

[1] Treiman [1959]; my italics.

[2] Nagel [1961], pp. 65–6.

[3] Cf. *above*, p. 119, footnote 1.

experiments' based on measurements of light velocity which 'disproved' the corpuscular theory and which turned out later to be erroneous in the light of relativity theory. These 'crucial experiments' were later deleted from the justificationist textbooks as manifestations of shameful short-sightedness or even of envy. (Recently they reappeared in some new textbooks, this time to illustrate the inescapable irrationality of scientific fashions.) However, in those cases in which ostensibly 'crucial experiments' were indeed *later* borne out by the defeat of the programme, historians charged those who resisted them with stupidity, jealousy, or unjustified adulation of the father of the research programme in question. (Fashionable 'sociologists of knowledge'—or 'psychologists of knowledge'—tend to explain positions in purely social or psychological terms when, as a matter of fact, they are determined by rationality principles. A typical example is the explanation of Einstein's opposition to Bohr's complementarity principle on the ground that 'in 1926 Einstein was forty-seven years old. Forty-seven may be the prime of life, but not for physicists'.[1])

In the light of this paper, the utopian idea of instant rationality becomes a hallmark of most brands of epistemology. Justificationists wanted scientific theories to be proved even before they were published; probabilists hoped a machine could flash up instantly the value (degree of confirmation) of a theory, given the evidence; naive falsificationists hoped that elimination at least was the instant result of the verdict of *experiment*.[2] I hope I have shown that *all these theories of instant rationality—and instant learning—fail*. The case studies of this section show that rationality works much slower than most people tend to think, and, even then, fallibly. Minerva's owl flies at dusk. I also hope I have shown that the *continuity* in science, the *tenacity* of some theories, the rationality of a certain

[1] Bernstein [1961], p. 129. In order to appraise progressive and degenerating elements in rival problemshifts one must understand the *ideas* involved. But the sociology of knowledge frequently serves as a successful cover for illiteracy: most sociologists of knowledge do not understand—or even care for—the ideas; they watch the socio-psychological patterns of behaviour. Popper used to tell a story about a 'social psychologist', Dr. X, studying scientists' group behaviour. He went into a physics seminar to study the psychology of science. He observed the 'emergence of a leader', the 'rallying round effect' in some and the 'defence-reaction' in others, the correlation between age, sex and aggressive behaviour, etc. (Dr. X claimed to have used some sophisticated small-sample techniques of modern statistics.) At the end of the enthusiastic account Popper asked Dr. X: 'What was the *problem* the group was discussing?' Dr. X was surprised: 'Why do you ask? I did not listen to the *words*! Anyway, what has *that* to do with the psychology of knowledge?'

[2] Of course, naive falsificationists may take some time to reach the 'verdict of experiment': the experiment has to be repeated and critically considered. But once the discussion ends up in an agreement among the experts, and thus a 'basic statement' becomes 'accepted', and it has been decided which specific theory was hit by it, the naive falsificationist will have little patience with those who still 'prevaricate'.

amount of dogmatism, can only be explained if we construe science as a battleground of research programmes rather than of isolated theories. One can understand very little of the growth of science when our paradigm of a chunk of scientific knowledge is an isolated theory like 'All swans are white', standing aloof, without being embedded in a major research programme. *My account implies a new criterion of demarcation between 'mature science', consisting of research programmes, and 'immature science', consisting of a mere patched up pattern of trial and error.*[1] For instance, we may have a conjecture, have it refuted and then rescued by an auxiliary hypothesis which is not *ad hoc* in the senses which we had earlier discussed. It may predict novel facts some of which may even be corroborated.[2] Yet one may achieve such 'progress' with a patched up, arbitrary series of disconnected theories. Good scientists will not find such makeshift progress satisfactory; they may even reject it as not genuinely scientific. They will call such auxiliary hypotheses merely 'formal', 'arbitrary', 'empirical', 'semi-empirical', or even '*ad hoc*'.[3]

Mature science consists of research programmes in which not only novel facts but, in an important sense, also novel auxiliary theories, are anticipated; mature science—unlike pedestrian trial-and-error—has 'heuristic power'. Let us remember that in the positive heuristic of a powerful programme there is, right at the start, a general outline of how to build the protective belts: this heuristic power generates *the autonomy of theoretical science.*[4]

This *requirement of continuous growth* is my rational reconstruction of the widely acknowledged requirement of 'unity' or 'beauty' of science. It highlights the weakness of *two*—apparently very different—types of theorizing. First, it shows up the weakness of programmes which, like Marxism or Freudism, are, no doubt, 'unified', which give a major sketch of the sort of auxiliary theories they are going to use in absorbing anomalies, but which

[1] The elaboration of this demarcation in the two following paragraphs was improved in the press, following invaluable discussions with Paul Meehl in Minneapolis in 1969.

[2] Earlier I distinguished, following Popper, two criteria of *adhocness*. I called *ad hoc*$_1$ theories which had no excess content over their predecessors (or competitors) that is, which did not predict any *novel* facts; I called *ad hoc*$_2$ theories which predicted novel facts but completely failed: none of their excess content got corroborated (cf. *above*, p. 124, footnote 3, and p. 125, footnote 1).

[3] Planck's radiation formula—given in his [1900*a*]—is a good example: cf. *above*, p. 167, footnote 2. We may call such hypotheses which are not *ad hoc*$_1$, not *ad hoc*$_2$, but still unsatisfactory in the sense specified in the text, *ad hoc*$_3$. These three—unfailingly pejorative—usages of *ad hoc* may provide a satisfactory entry in the *Oxford English Dictionary*.

It is intriguing to note that 'empirical' and 'formal' are both used as synonyms for our *ad hoc*$_3$.

Meehl, in his brilliant [1967], reports that in contemporary psychology—especially in social psychology—many alleged 'research programmes' in fact consist of chains of such *ad hoc*$_3$ stratagems.

[4] Cf. *above*, p. 137.

unfailingly devise their actual auxiliary theories in the wake of facts without, at the same time, anticipating others. (What *novel* fact has Marxism *predicted* since, say, 1917?) Secondly, it hits patched-up, unimaginative series of pedestrian 'empirical' adjustments which are so frequent, for instance, in modern social psychology. Such adjustments may, with the help of so-called 'statistical techniques', make some 'novel' predictions and may even conjure up some irrelevant grains of truth in them. But this theorizing has no unifying idea, no heuristic power, no continuity. They do not add up to a genuine research programme and are, on the whole, worthless.[1]

My account of scientific rationality, although based on Popper's, leads away from some of his general ideas. I endorse to some extent both Le Roy's conventionalism with regard to theories and Popper's conventionalism with regard to basic propositions. In this view scientists (and as I have shown, mathematicians too[2]) are not irrational when they tend to ignore counterexamples or as they prefer to call them, 'recalcitrant' or 'residual' instances, and follow the sequence of problems as prescribed by the positive heuristic of their programme, and elaborate—and apply—their theories regardless.[3] Contrary to Popper's falsificationist morality, scientists frequently and *rationally* claim 'that the experimental results are not reliable, or that the discrepancies which are asserted to exist between the experimental results and the theory are only apparent and

[1] After reading Meehl [1967] and Lykken [1968] one wonders whether the function of statistical techniques in the social sciences is not primarily to provide a machinery for producing phoney corroborations and thereby a semblance of 'scientific progress' where, in fact, there is nothing but an increase in pseudo-intellectual garbage. Meehl writes that 'in the physical sciences, the usual result of an improvement in experimental design, instrumentation, or numerical mass of data, is to increase the difficulty of the "observational hurdle" which the physical theory of interest must successfully surmount; whereas, in psychology and some of the allied behaviour sciences, the usual effect of such improvement in experimental precision is to provide an easier hurdle for the theory to surmount'. Or, as Lykken put it: 'Statistical significance [in psychology] is perhaps the least important attribute of a good experiment; it is never a sufficient condition for claiming that a theory has been usefully corroborated, that a meaningful empirical fact has been established, or that an experimental report ought to be published.' It seems to me that most theorizing condemned by Meehl and Lykken may be *ad hoc*$_3$. Thus the methodology of research programmes might help us in devising laws for stemming this intellectual pollution which may destroy our cultural environment even earlier than industrial and traffic pollution destroys our physical environment.

[2] Cf. my [1963–4].

[3] Thus the *methodological* asymmetry between universal and singular statements vanishes. We may adopt either by convention: in the 'hard core' we decide to 'accept' universal, in the 'empirical basis' singular, statements. The *logical* asymmetry between universal and singular statements is fatal only for the dogmatic inductivist who wants to learn only from hard experience and logic. The conventionalist can, of course, 'accept' this *logical* asymmetry: he does not have to be (although he *may* be) also an inductivist. He 'accepts' some universal statements, but not because he claims to deduce (or induce) them from singular ones.

that they will disappear with the advance of our understanding'.[1] When doing so, they may *not* be 'adopting the very reverse of that critical attitude which . . . is the proper one for the scientist'.[2] Indeed, Popper is right in stressing that 'the dogmatic attitude of sticking to a theory as long as possible is of considerable significance. Without it we could never find out what is in a theory—we should give the theory up before we had a real opportunity of finding out its strength; and in consequence no theory would ever be able to play its role of bringing order into the world, of preparing us for future events, of drawing our attention to events we should otherwise never observe'.[3] Thus the 'dogmatism' of 'normal science' does not prevent growth as long as we combine it with the Popperian recognition that there is good, progressive normal science and that there is bad, degenerating normal science, and as long as we retain the *determination* to eliminate, under certain objectively defined conditions, some research programmes.

The dogmatic attitude in science—which would explain its stable periods—was described by Kuhn as a prime feature of 'normal science'.[4] But Kuhn's conceptual framework for dealing with continuity in science is socio-psychological: mine is normative. I look at continuity in science through 'Popperian spectacles'. Where Kuhn sees 'paradigms', I *also* see rational 'research programmes'.

4. THE POPPERIAN VERSUS THE KUHNIAN RESEARCH PROGRAMME

Let us now sum up the Kuhn–Popper controversy.

We have shown that Kuhn is right in objecting to naive falsificationism, and also in stressing the *continuity* of scientific growth, the *tenacity* of some scientific theories. But Kuhn is wrong in thinking that by discarding naive falsificationism he has discarded thereby all brands of falsificationism. Kuhn objects to the entire Popperian research programme, and he excludes *any* possibility of a rational reconstruction of the growth of science. In a

[1] Popper [1934], section 9.

[2] *Ibid.*

[3] Popper [1940], first footnote. We find a similar remark in his [1963], p. 49. But these remarks are in *prima facie* contradiction with some of his remarks in [1934] (quoted *above*, p. 111), and therefore may only be interpreted as signs of a growing awareness by Popper of an undigested anomaly in his own research programme.

[4] Indeed, my demarcation criterion between mature and immature science can be interpreted as a Popperian absorption of Kuhn's idea of 'normality' as a hallmark of [mature] science; and it also improved on our earlier argument against regarding such highly falsifiable statements as scientific. (Cf. *above*. p. 102.)

Incidentally, this demarcation between mature and immature science appears already in my [1961] and [1963–4], where I called the former 'deductive guessing' and the latter 'naive trial and error'. (See e.g. [1963–4], section 7(c): 'Deductive guessing versus naive guessing.')

succint comparison of Hume, Carnap and Popper, Watkins points out that the growth of science is inductive and irrational according to Hume, inductive and rational according to Carnap, non-inductive and rational according to Popper.[1] But Watkins's comparison can be extended by adding that it is non-inductive and irrational according to Kuhn. *In Kuhn's view there can be no logic, but only psychology of discovery.*[2] For instance, in Kuhn's conception, anomalies, inconsistencies *always* abound in science, but in 'normal' periods the dominant paradigm secures a pattern of growth which is eventually overthrown by a 'crisis'. There is no particular rational cause for the appearance of a Kuhnian 'crisis'. 'Crisis' is a psychological concept; it is a contagious panic. Then a new 'paradigm' emerges, incommensurable with its predecessor. There are no rational standards for their comparison. Each paradigm contains its own standards. The crisis sweeps away not only the old theories and rules but also the standards which made us respect them. The new paradigm brings a totally new rationality. There are no super-paradigmatic standards. The change is a bandwagon effect. Thus *in Kuhn's view scientific revolution is irrational, a matter for mob psychology.*

The reduction of philosophy of science to psychology of science did not start with Kuhn. An earlier wave of 'psychologism' followed the breakdown of justificationism. For many, justificationism represented the only possible form of rationality: the end of justificationism meant the end of rationality. The collapse of the thesis that scientific theories are provable, that the progress of science is cumulative, made justificationists panic. If 'to discover is to prove', but nothing is provable, then there can be no discoveries, only discovery-claims. Thus disappointed justificationists—ex-justificationists—thought that the elaboration of rational standards was a hopeless enterprise and that all one can do is to study—and imitate—the Scientific Mind, as it is exemplified in famous scientists. After the collapse of Newtonian physics, Popper elaborated new, non-justificationist critical standards. Now some of those who had already learned of the collapse of justificationist rationality now learned, mostly by hearsay, of Popper's colourful slogans which suggested naive falsificationism. Finding them untenable, they identified the collapse of naive falsificationism with the end of rationality itself. The elaboration of rational standards was again regarded as a hopeless enterprise; the best one can do is to study, they thought once again, the Scientific Mind.[3] Critical philosophy was to be replaced by what Polanyi called a 'post-critical' philosophy. But the Kuhnian research

[1] Watkins [1968], p. 281.

[2] Kuhn [1965]. But this position is already implicit in his [1962].

[3] Incidentally, just as some earlier ex-justificationists led the wave of sceptical irrationalism, so now some ex-falsificationists lead the *new* wave of sceptical irrationalism and anarchism. This is best exemplified in Feyerabend [1970].

programme contains a new feature: we have to study not the mind of the individual scientist but the mind of the Scientific Community. Individual psychology is now replaced by social psychology; imitation of the great scientists by submission to the collective wisdom of the community.

But Kuhn overlooked Popper's sophisticated falsificationism and the research programme he initiated. Popper replaced the central problem of classical rationality, *the old problem of foundations*, with *the new problem of fallible-critical growth*, and started to elaborate objective standards of this growth. In this paper I have tried to develop his programme a step further. I think this small development is sufficient to escape Kuhn's strictures.[1]

The reconstruction of scientific progress as proliferation of rival research programmes and progressive and degenerative problemshifts gives a picture of the scientific enterprise which is in many ways different from the picture provided by its reconstruction as a succession of bold theories and their dramatic overthrows. Its main aspects were developed from Popper's ideas and, in particular, from his ban on 'conventionalist', that is, content-decreasing, stratagems. The main difference from Popper's original version is, I think, that in my conception criticism does not—and must not—kill as fast as Popper imagined. *Purely negative, destructive criticism, like 'refutation' or demonstration of an inconsistency does not eliminate a programme. Criticism of a programme is a long and often frustrating process and one must treat budding programmes leniently.*[2] One may, of course, show up the degeneration of a research programme, but it is only *constructive criticism* which, with the help of rival research programmes, can achieve real successes; and dramatic spectacular results become visible only with hindsight and rational reconstruction.

Kuhn certainly showed that the psychology of science can reveal important and, indeed, sad truths. But the psychology of science is not autonomous; for *the—rationally reconstructed—growth of science takes place*

[1] Indeed, as I had already mentioned, *my concept of a 'research programme' may be construed as an objective, 'third world' reconstruction of Kuhn's socio-psychological concept of paradigm*': thus the Kuhnian 'Gestalt-switch' can be performed without removing one's Popperian spectacles.

(I have not dealt with Kuhn's and Feyerabend's claim that theories cannot be eliminated on any *objective* grounds because of the 'incommensurability' of rival theories. Incommensurable theories are neither inconsistent with each other, nor comparable for content. But we can *make* them, by a dictionary, inconsistent and their content comparable. If we want to eliminate a programme, we need some methodological determination. This determination is the heart of methodological falsificationism; for instance, no result of statistical sampling is ever inconsistent with a statistical theory unless we *make them* inconsistent with the help of Popperian rejection rules, cf. *above*, p. 109.)

[2] The reluctance of economists and other social scientists to accept Popper's methodology may have been partly due to the destructive effect of naive falsificationism on budding research programmes.

essentially in the world of ideas, in Plato's and Popper's 'third world', in the world of articulated knowledge which is independent of knowing subjects.[1] *Popper's research programme* aims at a description of this objective scientific *growth*.[2] *Kuhn's research programme* seems to aim at a description of *change* in the ('normal') scientific mind (whether individual or communal).[3] But the mirror-image of the third world in the mind of the individual—even in the mind of the 'normal'—scientists is usually a caricature of the original; and to describe this caricature without relating it to the third-world original might well result in a caricature of a caricature. One cannot understand the history of science without taking into account the interaction of the three worlds.

APPENDIX

POPPER, FALSIFICATIONISM AND THE 'DUHEM-QUINE THESIS'

Popper began as a dogmatic falsificationist in the 1920s; but he soon realized the untenability of this position and published nothing before he invented *methodological falsificationism*. This was an entirely new idea in the philosophy of science and it clearly originates with Popper, who put it forward as a solution to the difficulties of dogmatic falsificationism. Indeed, the conflict between the theses that science is both critical and

[1] The *first* world is the material world, the *second* is the world of consciousness, the *third* is the world of propositions, truth, standards: the world of objective knowledge. The modern *loci classici* on this subject are Popper [1968*a*] and Popper [1968*b*]; also, cf. Toulmin's impressive programme set out in his [1967]. It should be mentioned here that many passages of Popper [1934] and even of [1963] sound like descriptions of a psychological contrast between the Critical Mind and the Inductivist Mind. But Popper's psychologistic terms can be, to a large extent, reinterpreted in third-world terms. Cf. Musgrave [1971].

[2] In fact, Popper's research programme extends beyond science. The concepts of 'progressive' and 'degenerating' problemshifts, the idea of proliferation of theories can be generalized to any sort of rational discussion and thus serve as tools for a general theory of criticism; cf. my [1970].

[3] *Actual* state of minds, beliefs, etc., belong to the second world; states of the *normal* mind belong to a limbo between the second and third. The study of actual scientific minds belongs to *psychology*; the study of the 'normal' (or 'healthy' etc.) mind belongs to a *psychologistic philosophy of science*. There are *two kinds of psychologistic philosophies of science*. According to one kind there can be no philosophy of science: only a psychology of individual scientists. According to the other kind there is a psychology of the 'scientific', 'ideal' or 'normal' mind: this turns philosophy of science into a psychology of this ideal mind and, in addition, offers a psychotherapy for turning one's mind into an ideal one. I discuss this second kind of psychologism in detail in my [1970]. Kuhn does not seem to have noticed this distinction.

fallible is one of the central problems in Popperian philosophy. While Popper offered a coherent formulation and criticism of dogmatic falsificationism, he never made a sharp distinction between naive and sophisticated falsificationism. In an earlier paper,[1] I distinguished three Poppers: $Popper_0$, $Popper_1$ and $Popper_2$. $Popper_0$ is the dogmatic falsificationist who never published a word: he was invented—and 'criticized'—first by Ayer and then by many others.[2] This paper will, I hope, finally kill this ghost. $Popper_1$ is the naive falsificationist, $Popper_2$ the sophisticated falsificationist. The *real* Popper developed from dogmatic to a naive version of methodological falsificationism in the twenties; he arrived at the *'acceptance rules' of sophisticated falsificationism* in the fifties. The transition was marked by his adding to the original requirement of testability the 'second' requirement of 'independent testability',[3] and then the 'third' requirement that some of these independent tests should result in corroborations.[4] But the real Popper never abandoned his earlier (naive) *falsification rules*. He has demanded, until this day, that *'criteria of refutation* have to be laid down beforehand: it must be agreed, which observable situations, if actually observed, mean that the theory is refuted'.[5] He still construes 'falsification' as the result of a duel between theory and observation, without another, better theory *necessarily* being involved. The real Popper has never explained in detail the appeal procedure by which some 'accepted basic statements' may be eliminated. Thus the real Popper consists of $Popper_1$ together with some elements of $Popper_2$.

The idea of a demarcation between progressive and degenerating problemshifts, as discussed in this paper, is based on Popper's work:

[1] Cf. my [1968*b*].

[2] Ayer seems to have been the first to attribute dogmatic falsificationism to Popper. (Ayer also invented the myth that according to Popper 'definite confutability' was a criterion not only of the empirical but also of the meaningful character of a proposition: cf. his [1936], chapter 1, p. 38 of the second edition.) Even today, many philosophers (cf. Juhos [1966] or Nagel [1967]) criticize the strawman $Popper_0$. Medawar, in his [1967], called *dogmatic* falsificationism 'one of the strongest ideas' in Popper's methodology. Nagel, reviewing Medawar's book, criticized Medawar for 'endorsing' what he too believes to be 'Popper's claims' (Nagel [1967], p. 70). Nagel's criticism convinced Medawar that 'the act of falsification is not immune to human error' (Medawar [1969], p. 54). But Medawar and Nagel misread Popper: his *Logik der Forschung* is the strongest ever criticism of dogmatic falsificationism.

One may take a charitable view of Medawar's mistake: for brilliant scientists whose speculative talent was thwarted under the tyranny of an inductivist logic of discovery, falsificationism, even in its dogmatic form, was bound to have a tremendous liberating effect. (Besides Medawar, another Nobel Prize winner, Eccles, learned from Popper to replace his original caution by bold falsifiable speculation: cf. his [1964], pp. 274–5.)

[3] Popper [1957].

[4] Popper [1963], pp. 242 ff.

[5] Popper [1963], p. 38, footnote 3.

indeed this demarcation is almost identical with his celebrated demarcation criterion between science and metaphysics.[1]

Popper originally had only the *theoretical* aspect of problemshifts in mind, which is hinted at in section 20 of his [1934] and developed in his [1957].[2] He added a discussion of the *empirical* aspect of problemshifts only later, in his [1963].[3] However, Popper's ban on 'conventionalist stratagems' is in some respects too strong, in others too weak. It is too *strong*, for, according to Popper, a new version of a progressive programme *never* adopts a content-decreasing stratagem to absorb an anomaly, it *never* says things like 'all bodies are Newtonian except for seventeen anomalous ones'. But since unexplained anomalies always abound, I allow such formulations; an explanation is a step forward (that is, 'scientific') if it explains at least *some* previous anomalies which were not explained 'scientifically' by its predecessor. As long as anomalies are regarded as genuine (though not necessarily urgent) problems, it does not matter much whether we dramatize them as 'refutations' or de-dramatize them as 'exceptions': the difference *then* is only a linguistic one. (This degree of tolerance of *ad hoc* stratagems allows us to progress even on inconsistent foundations. Problemshifts may then be progressive in spite of inconsistencies.[4]) However, Popper's ban on content-decreasing stratagems is also too *weak*: it cannot deal, for instance, with the 'tacking paradox',[5] and does not ban *ad hoc*$_3$ stratagems.[6] These can be eliminated only by the requirement that *the auxiliary hypotheses should be formed in accordance with the positive heuristic of a genuine research programme.* This new requirement brings us to the problem of *continuity in science.*

[1] If the reader is in doubt about the authenticity of my reformulation of Popper's demarcation criterion, he should re-read the relevant parts of Popper [1934] with Musgrave [1968] as a guide. Musgrave wrote his [1968] against Bartley who, in his [1968], mistakenly attributed to Popper the demarcation criterion of naive falsificationism, as formulated *above*, p. 109.

[2] In his [1934], Popper was primarily concerned with a ban on *surreptitious ad hoc* adjustments. Popper (Popper$_1$) demands that the design of a potentially negative crucial experiment must be presented together with the theory, and then the verdict of the experimental jury humbly accepted. It follows that conventionalist stratagems, which *after* the verdict give a retrospective twist to the original theory in order to escape the verdict, are *eo ipso* ruled out. But if we admit the refutation and *then* reformulate the theory with the help of an *ad hoc* stratagem, we may admit it as a '*new*' theory; and if it is testable, then Popper$_1$ accepts it for new criticism: 'Whenever we find that a system has been rescued by a conventionalist stratagem, we shall test it afresh, and reject it, as circumstances may require' (Popper [1934], section 20).

[3] For details, cf. my [1968*a*], especially pp. 388–90.

[4] Cf. *above*, pp. 142 ff. This tolerance is rarely, if ever, found in textbooks of scientific method.

[5] Cf. *above*, p. 131.

[6] Cf. *above*, p. 175, footnote 3.

The problem of *continuity* in science was raised by Popper and his followers long ago. When I proposed my theory of growth based on the idea of competing research programmes, I again followed, and tried to improve, Popperian tradition. Popper himself, in his [1934], had already stressed the heuristic importance of 'influential metaphysics',[1] and was regarded by some members of the Vienna Circle as a champion of dangerous metaphysics.[2] When his interest in the role of metaphysics revived in the 1950s, he wrote a most interesting 'Metaphysical Epilogue' about 'metaphysical research programmes' to his *Postscript: After Twenty Years* —in galleys since 1957.[3] But Popper associated tenacity not with *methodological irrefutability* but rather with *syntactical irrefutability.* By 'metaphysics' he meant syntactically specifiable statements like 'all-some' statements and purely existential statements. No basic statements could conflict with them because of their logical form. For instance, 'for all metals there is a solvent' would, in this sense, be 'metaphysical', while Newton's theory of gravitation, taken in isolation, would not be.[4] Popper, in the 1950s, also raised the problem of how to criticize metaphysical theories and suggested

[1] Cf. e.g. his [1934], end of section 4; also cf. his [1968*c*], p. 93. One should remember that such importance was denied to metaphysics by Comte and Duhem. The people who did most to reverse the anti-metaphysical tide in the philosophy and the historiography of science were Burtt, Popper and Koyré.

[2] Carnap and Hempel were trying, in their reviews of the book, to defend Popper against this charge (cf. Carnap [1935] and Hempel [1937]). Hempel wrote: '[Popper] stresses strongly certain features of his approach which are common with the approach of somewhat metaphysically oriented thinkers. It is to be hoped that this valuable work will not be misinterpreted as if it meant to allow for a new, perhaps even logically defensible, metaphysics.'

[3] A passage of this *Postscript* is here worth quoting: 'Atomism is an . . . excellent example of a non-testable metaphysical theory whose influence upon science exceeded that of many testable theories . . . The latest and greatest so far was the programme of Faraday, Maxwell, Einstein, de Broglie, and Schrödinger, of conceiving the world . . . in terms of continuous fields . . . Each of these metaphysical theories functioned, long before it became testable, as a programme for science. It indicated the direction in which satisfactory explanatory theories of science may be found, and it made possible something like an appraisal of the depth of a theory. In biology, the theory of evolution, the theory of the cell, and the theory of bacterial infection, have all played similar parts, at least for a time. In psychology, sensualism, atomism (that is, the theory that all experiences are composed of last elements, such as, for example, sense data) and psycho-analysis should be mentioned as metaphysical research programmes . . . Even purely existential assertions have sometimes proved suggestive and even fruitful in the history of science even if they never became part of it. Indeed, few metaphysical theories exerted a greater influence upon the development of science than the purely metaphysical one: "There exists a substance which can turn base metals into gold (that is, a philosopher's stone)", although it is non-falsifiable, was never verified, and is now believed by nobody.'

[4] Cf. especially Popper [1934], section 66. In the 1959 edition he added a clarifying footnote (footnote *2) in order to stress that in *metaphysical* "all-some' statements the existential quantifier must be interpreted as 'unbounded'; but, of course, he had made this absolutely clear already in section 15 of the original text.

solutions.[1] Agassi and Watkins published several interesting papers on the role of this sort of 'metaphysics' in science, which all connected 'metaphysics' with the continuity of scientific progress.[2] My treatment differs from theirs first because I go much further than they in blurring the demarcation between [Popper's] 'science' and [Popper's] 'metaphysics': I do not even use the term 'metaphysical' any more. I only talk about *scientific* research programmes whose hard core is irrefutable, not because of syntactical but because of methodological reasons which have nothing to do with logical form. Secondly, separating sharply the *descriptive problem* of the psychologico-historical role of metaphysics from the *normative problem* of how to distinguish progressive from degenerating research programmes, I elaborate the latter problem further than they had done.

Finally, I should like to discuss the '*Duhem-Quine thesis*', and its relation to falsificationism.[3]

According to the 'Duhem-Quine thesis', given sufficient imagination, any theory (whether consisting of one proposition or of a finite conjunction of many) can be permanently saved from 'refutation' by some suitable adjustment in the background knowledge in which it is embedded. As Quine put it: 'Any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system . . . Conversely, by the same token, no statement is immune to revision.'[4] Moreover, the 'system' is nothing less than 'the whole of science'. 'A recalcitrant experience can be accommodated by any of various alternative reëvaluations in various alternative quarters of the total system [including the possibility of reëvaluating the recalcitrant experience itself].'[5]

This thesis has two very different interpretations. In its *weak interpretation* it only asserts the impossibility of a direct experimental hit on a narrowly specified theoretical target and the logical possibility of shaping science in indefinitely many different ways. The weak interpretation hits only dogmatic, not methodological, falsificationism: it only denies the possibility of a *disproof* of any *separate* component of a theoretical system.

In its *strong interpretation* the Duhem–Quine thesis excludes any *rational* selection rule among the alternatives; this version is inconsistent with all forms of methodological falsificationism. The two interpretations have not been clearly separated, although the difference is methodologically vital. Duhem seems to have held only the weak interpretation: for

[1] Cf. especially his [1963], pp. 198–9 (first published in 1958).
[2] Cf. Watkins [1957] and [1958] and Agassi [1962] and [1964].
[3] This concluding part of the *Appendix* was added in print.
[4] Quine [1953], chapter ii. [5] *Ibid.* The clause in the square brackets is mine.

him the selection is a matter of 'sagacity': we must always make the right choices in order to get nearer to 'natural classification'.[1] On the other hand, Quine, in the tradition of the American pragmatism of James and Lewis, seems to hold a position very near to the strong interpretation.[2]

Let us now have a closer look at the weak Duhem–Quine thesis. Let us take a 'recalcitrant experience' expressed in an 'observation statement' O' which is inconsistent with a conjunction of theoretical (and 'observational') statements $h_1, h_2 \ldots h_n, I_1, I_2 \ldots I_n$, where h_i are theories and I_i the corresponding initial conditions. In the 'deductive model', $h_1 \ldots h_n, I_1 \ldots I_n$ logically imply O; but O' is observed which implies *not-O*. Let us also assume that the premisses are independent and are all necessary for deducing O.

In this case we may restore consistency by altering *any* of the sentences in our deductive model. For instance, let h_1 be: 'whenever a thread is loaded with a weight exceeding that which characterizes the tensile strength of the thread, then it will break'; let h_2 be: 'the weight characteristic for this thread is 1 *lb.*'; let h_3 be: 'the weight put on this thread was 2 *lbs*'. Let, finally, h_0 be: 'an iron weight of 2 *lbs* was put on the thread located in the space-time position P and it did not break'. One may solve the problem in many ways. To give a few examples: (1) We reject h_1; we replace the expression 'is loaded with a weight' by 'is pulled by a force'; we introduce a new initial condition: there was a hidden magnet (or hitherto unknown force) located in the laboratory ceiling. (2) We reject h_2; we propose that the tensile strength *does* depend on how moist threads are; the tensile strength of the actual thread, since it got moist, was 2 *lbs*. (3) We reject h_3; the weight was only 1 *lb*; the scales went wrong. (4) We reject h_0; the thread did not break; it was only *observed* to break, but the professor who proposed h_1 & h_2 & h_3 was a well-known bourgeois liberal and his revolutionary laboratory assistants consistently *saw* his hypotheses refuted when in fact they were confirmed. (5) We reject h_3; the thread was not a 'thread', but a 'superthread', and 'superthreads' never break.[3] We could go on

[1] An experiment, for Duhem, can never *alone* condemn an isolated theory (such as the hard core of a research programme): for such 'condemnation' we *also* need 'common sense', 'sagacity', and, indeed, good metaphysical instinct which leads us towards (or *to*) 'a certain supremely eminent order'. (See the end of the *Appendix* of the second edition of his [1906].)
[2] Quine speaks of statements having 'varying distances from a sensory periphery', and thus more or less exposed to change. But both the sensory periphery and the metric are hard to define. According to Quine 'the considerations which guide [man] in warping his scientific heritage to fit his continuing sensory peripheries are, where rational, pragmatic' (Quine [1953]). But 'pragmatism' for Quine, as for James or LeRoy, is only psychological comfort; and I find it irrational to call this 'rational'.
[3] For such 'concept-narrowing defences' and 'concept-stretching refutations', cf. my [1963–64].

indefinitely. Indeed, there are infinitely many possibilities of how to replace—given sufficient imagination—any of the premisses (*in the deductive model*) by invoking a change in some *distant* part of our total knowledge (*outside the deductive model*) and thereby restore consistency.

Can we formulate this trivial observation by saying that '*each test is a challenge to the whole of our knowledge*'? I do not see any reason why not. The resistance of some falsificationists to this 'holistic dogma of the "global" character of all tests'[1] is due only to a semantic conflation of two different notions of 'test' (or 'challenge') which a recalcitrant experimental result presents to our knowledge.

The Popperian interpretation of a 'test' (or 'challenge') is that the result (O) contradicts ('challenges') a finite, well-specified conjunction of premisses (T): O & T cannot be true. But no protagonist of the Duhem–Quine argument would deny this point.

The Quinean interpretation of 'test' (or 'challenge') is that the *replacement* of O & T may invoke some change also outside O and T. The successor to O & T may be inconsistent with some H in some distant part of knowledge. But no Popperian would deny this point.

The conflation of the two notions of testing led to some misunderstandings and logical blunders. Some people felt intuitively that the *modus tollens* from refutation may 'hit' very distant premisses in our total knowledge and therefore were trapped in the idea that the '*ceteris paribus* clause' is a premiss which is joined *conjunctively* with the obvious premisses. But this 'hit' is achieved not by *modus tollens* but as a result of our subsequent replacement of our original deductive model.[2]

Thus 'Quine's weak thesis' trivially holds. But 'Quine's strong thesis' will be strenuously opposed, both by the naive and the sophisticated falsificationist.

The naive falsificationist insists that if we have an inconsistent set of scientific statements, we first must select from among them (1) a theory under test (to serve as a *nut*); then we must select (2) an accepted basic statement (to serve as a *hammer*) and the rest will be uncontested background knowledge (to provide an *anvil*). And in order to put teeth into this position, we must offer a method of 'hardening' the 'hammer' and the 'anvil' in order to enable us to crack the 'nut', and thus perform a 'negative

[1] Popper [1963], chapter 10, section xvi.

[2] The *locus classicus* of this confusion is Canfield's and Lehrer's wrongheaded criticism of Popper in their [1961]; Stegmüller followed them into the logical morass ([1966], p. 7). Coffa contributed to the clarification of the issue ([1968]).

Unfortunately, my own phraseology in this paper in places suggests that the '*ceteris paribus* clause' is an independent premiss in the theory under test. My attention was drawn to this easily repairable defect by Colin Howson.

crucial experiment'. But naive 'guessing' of this division is too arbitrary, it does not give us any serious hardening. (Grünbaum, on the other hand, swallowing his falsificationist pride, stoops down to accept help from the inductivist Salmon. He now applies Salmon's Reichenbachian theory of probability of hypotheses in order to show that, at least in some sense, the 'hammer' and the 'anvil' have high posterior probabilities and therefore are 'hard' enough for being used as a nutcracker.[1])

The sophisticated falsificationist allows *any* part of the body of science to be replaced *but* only on the condition that it is replaced in a 'progressive' way, so that the replacement successfully anticipates novel facts. In his rational reconstruction of falsification 'negative crucial experiments' play no role. He sees nothing wrong with a group of brilliant scientists conspiring to pack everything they can into their favourite research programme ('conceptual framework', if you wish) with a sacred hard core. As long as their genius—and luck—enables them to expand their programme '*progressively*', while sticking to its hard core, they are allowed to do it. And if a genius comes determined to *replace* ('progressively') a most uncontested and corroborated theory which he happens to dislike on philosophical, aesthetic or personal grounds, good luck to him. If two teams, pursuing rival research programmes, compete, the one with more creative talent is likely to succeed—unless God punishes them with an extreme lack of empirical success. The direction of science is determined primarily by human creative imagination and not by the universe of facts which surrounds us. Creative imagination is likely to find corroborating novel evidence even for the most 'absurd' programme, if the search has sufficient drive.[2] This look-out for *new confirming evidence* is perfectly permissible. Scientists dream up phantasies and then pursue a highly selective hunt for new facts which fit these phantasies. This process may be described as

[1] Grünbaum [1969]. He previously took a position which was one of radical dogmatic falsificationism and claimed that we *can* ascertain the falsity of scientific hypotheses (e.g. Grünbaum [1959*b*] and [1960]). His concrete case studies were thought-provoking for the philosopher and challenging for the physicist. But after criticisms from Feyerabend (cf. his [1959]), Laudan (cf. his [1965]), and others, he had to modify his position: such falsification cannot always be 'ascertained irrevocably': 'At least in some cases, we can ascertain the falsity of a component hypothesis to all scientific intents and purposes, although we cannot falsify it beyond any and all possibility of subsequent rehabilitation.'

[2] A typical such example is Newton's principle of gravitational attraction according to which bodies attract each other instantly from immense distances. Huyghens described this idea as 'absurd', Leibnitz as 'occult', and the best scientists of the age 'wondered how [Newton] could have given himself all the trouble of making such a number of investigations and difficult calculations that had no other foundation than this very principle' (cf. Koyré [1965], pp. 117–18). I had argued earlier that it is not so that theoretical progress is the merit of the theoretician but empirical success is *merely* a matter of luck. If the theoretician is *more* imaginative, it is likelier that his theoretical programme will achieve at least *some* empirical success. Cf. my [1968*a*], pp. 387–90.

'science creating its own universe' (as long as one remembers that 'creating' here is used in a provocative-idiosyncratic sense). A brilliant school of scholars (backed by a rich society to finance a few well-planned tests) might succeed in pushing any fantastic programme ahead, or, alternatively, if so inclined, in overthrowing any arbitrarily chosen pillar of 'established knowledge'.

The *dogmatic* falsificationist will throw up his hands in horror at this approach. He will see the spectre of Bellarmino's instrumentalism arising from the rubble under which Newtonian success of 'proven science' had buried it. He will accuse the sophisticated falsificationist of building arbitrary Procrustean pigeon hole systems and forcing the facts into them. He may even brand it as a revival of the unholy irrationalist alliance of James's crude pragmatism and of Bergson's voluntarism, triumphantly vanquished by Russell and Stebbing.[1] But our sophisticated falsificationism combines 'instrumentalism' (or 'conventionalism') with a strong empiricist requirement, which neither medieval 'saviours of phenomena' like Bellarmino, nor pragmatists like Quine and Bergsonians like Le Roy, had appreciated: the Leibnitz–Whewell–Popper requirement that *the—well planned—building of pigeon holes must proceed much faster than the recording of facts which are to be housed in them.* As long as this requirement is met, it does not matter whether we stress the 'instrumental' aspect of imaginative research programmes for finding novel facts and for making trustworthy predictions, or whether we stress the putative growing Popperian 'verisimilitude' (that is, the estimated difference between the truth-content and falsity-content) of their successive versions.[2] Sophisticated falsificationism thus combines the best elements of voluntarism, pragmatism and of the realist theories of empirical growth.

The sophisticated falsificationist sides neither with Galileo nor with Cardinal Bellarmino. He does not side with Galileo, for he claims that our basic theories may all be equally absurd and unverisimilar for the divine mind; and he does not side with Bellarmino, unless the Cardinal were to agree that scientific theories may yet lead, in the long run, to ever more true and ever fewer false consequences and, *in this strictly technical sense*, may have increasing 'verisimilitude'.[3]

[1] Cf. Russell [1914], Russell [1946] and Stebbing [1914]. Russell, a justificationist, despised conventionalism: 'As will has gone up in the scale, knowledge has gone down. This is the most notable change that has come over the temper of philosophy in our age. It was prepared by Rousseau and Kant . . .' ([1946], p. 787). Popper, of course, got some of his inspiration from Kant and Bergson. (Cf. his [1934], sections 2 and 4.)

[2] Cf. Popper [1963], chapter 10.

[3] *'Verisimilitude'* has two distinct meanings which must not be conflated. First, it may be used to mean intuitive truthlikeness of the theory; in this sense, in my view, all scientific theories created by the human mind are equally unverisimilar and 'occult'. Secondly, it may be used to mean the set-theoretical difference between the true and false consequences of a theory which we can never know but certainly may guess. It was Popper who used 'verisimilitude' as a precise technical term to denote this difference ([1963], chapter 10). But his claim that this explication corresponds closely to *the* original meaning is mistaken and misleading. In the original prepopperian usage 'verisimilitude' could mean either *intuitive* truthlikeness or a naive proto-version of Popper's *empirical* truthlikeness. Popper gives interesting quotations for the latter ([1963], pp. 399 ff.) but none for the former. But Bellarmino might have agreed that Copernican theory had high 'verisimilitude' in Popper's technical sense but not that it had verisimilitude in the first, intuitive sense. Most 'instrumentalists' are 'realists' in the sense that they agree that the [Popperian] 'verisimilitude' of scientific theories is likely to be growing; but they are not 'realists' in the sense that they would agree that, for instance, the Einsteinian field approach is *intuitively* closer to the Blueprint of the Universe than the Newtonian action at a distance. *The 'aim of science' may then be increasing Popperian 'verisimilitude', but does not have to be also increasing classical verisimilitude.* The latter, as Popper himself said, is, unlike the former, a 'dangerously vague and metaphysical' idea ([1963], p. 231).

Popper's 'empirical verisimilitude' in a sense rehabilitates the idea of *cumulative growth* in science. But the driving force of cumulative growth in 'empirical verisimilitude' is revolutionary conflict in 'intuitive verisimilitude'.

When Popper was writing his 'Truth, rationality and the growth of knowledge', I had an uneasy feeling about his identification of the two concepts of verisimilitude. Indeed, it was I who asked him: 'Can we really speak about *better* correspondence? Are there such things as *degrees* of truth? Is it not dangerously misleading to talk as if Tarskian truth were located somewhere in a kind of metrical or at least topological space so that we can sensibly say of two theories—say an earlier theory t_1 and a later theory t_2, that t_2 has superseded t_1, or progressed beyond t_1, by approaching more closely to the truth than t_1?' (Popper [1963], p. 232). Popper rejected my vague misgivings. He felt—rightly—that he was proposing a very important new idea. But he was mistaken in believing that his new, technical conception of 'verisimilitude' completely absorbed the problems centred on the old *intuitive* 'verisimilitude'. Kuhn says: 'To say, for example, of a field theory that it "approaches more closely to the truth" than an older matter-and-force theory should mean, *unless words are being oddly used*, that the ultimate constituents of nature are more like fields than like matter and force' (*this volume, below*, p. 265; my italics). Indeed, Kuhn is right, except that words are *normally* 'oddly used'. I hope that this note may contribute to the clarification of the problem involved.

REFERENCES

Agassi [1959]: 'How are Facts Discovered?', *Impulse*, **3**, No. 10, pp. 2–4.

Agassi [1962]: 'The Confusion between Physics and Metaphysics in the Standard Histories of Sciences', in the *Proceedings of the Tenth International Congress of the History of Science*, 1964, **1**, pp. 231–8.

Agassi [1964]: 'Scientific Problems and Their Roots in Metaphysics', in Bunge (*ed.*): *The Critical Approach to Science and Philosophy*, 1964, pp. 189–211.

Agassi [1966]: 'Sensationalism', *Mind*, N.S. **75**, pp. 1–24.

Agassi [1968]: 'The Novelty of Popper's Philosophy of Science', *International Philosophical Quarterly*, **8**, pp. 442–63.

Agassi [1969]: 'Popper on Learning from Experience', in Rescher (*ed.*): *Studies in the Philosophy of Science*, 1969.

Ayer [1936]: *Language, Truth and Logic*, 1936; second edition 1946.

Bartley [1968]: 'Theories of Demarcation between Science and Metaphysics', in Lakatos and Musgrave (*eds*): *Problems in the Philosophy of Science*, 1968, pp. 40–64.

Becke and Sitte [1933]: 'Zur Theorie des β-Zerfalls', *Zeitschrift für Physik*, **86**, pp. 105–19.

Bernal [1965]: *Science in History*, third edition, 1965.

Bernstein [1961]: *A Comprehensible World: On Modern Science and its Origins*, 1961.
Bethe and Peierls [1934]: 'The "Neutrino"', *Nature*, **133**, p. 532.
Bohr [1913*a*]: 'On the Constitution of Atoms and Molecules', *Philosophical Magazine*, **26**, pp. 1–25, 476–502 and 857–75.
Bohr [1913*b*]: Letter to Rutherford, 6.3.1913; published in Bohr [1963], pp. xxxviii–ix.
Bohr [1913*c*]: 'The Spectra of Helium and Hydrogen', *Nature*, **92**, pp. 231–2.
Bohr [1922]: 'The Structure of the Atom', Nobel Lecture.
Bohr [1926]: Letter to *Nature*, **117**, p. 264.
Bohr [1930]: 'Chemistry and the Quantum Theory of Atomic Constitution', Faraday Lecture 1930, *Journal of the Chemical Society*, 1932/1, pp. 349–84.
Bohr [1933]: 'Light and Life', *Nature*, **131**, pp. 421–3 and 457–9.
Bohr [1936]: 'Conservation Laws in Quantum Theory', *Nature*, **138**, pp. 25–6.
Bohr [1949]: 'Discussion with Einstein on Epistemological Problems in Atomic Physics', in Schilpp (*ed.*): *Albert Einstein, Philosopher-Scientist*, 1949, **1**, pp. 201–41.
Bohr [1963]: *On the Constitution of Atoms and Molecules*, 1963.
Born [1948]: 'Max Karl Ernst Ludwig Planck', *Obituary Notices of Fellows of the Royal Society*, **6**, 161–80.
Born [1954]: 'The Statistical Interpretation of Quantum Mechanics', *Nobel Lecture* 1954.
Braithwaite [1938]: 'The Relevance of Psychology to Logic', *Aristotelian Society Supplementary Volumes*, **17**, pp. 19–41.
Braithwaite [1953]: *Scientific Explanation*, 1953.
Callendar [1914]: 'The Pressure of Radiation and Carnot's Principle,' *Nature*, **92**, p. 553.
Canfield and Lehrer [1961]: 'A Note on Prediction and Deduction', *Philosophy of Science*, 1961, **28**, pp. 204–8.
Carnap [1932–3]: 'Über Protokollsätze', *Erkenntnis*, **3**, pp. 215–28.
Carnap [1935]: Review of Popper's [1934], *Erkenntnis*, **5**, pp. 290–4.
Coffa [1968]: 'Deductive Predictions', *Philosophy of Science*, **35**, pp. 279–83.
Crookes [1886]: Presidential Address to the Chemistry Section of the British Association, *Report of British Association*, 1886, pp. 558–76.
Crookes [1888]: Report at the Annual General Meeting, *Journal of the Chemical Society*, **53**, pp. 487–504.
Davisson [1937]: 'The Discovery of Electron Waves', *Nobel Lecture*, 1937.
Dirac [1936]: 'Does Conservation of Energy Hold in Atomic Processes?', *Nature*, **137**, pp. 298–9.
Dirac [1951]: 'Is there an Aether?', *Nature*, **168**, pp. 906–7.
Dorling [1968]: 'Length Contraction and Clock Synchronisation: The Empirical Equivalence of the Einsteinian and Lorentzian Theories', *The British Journal for the Philosophy of Science*, **19**, pp. 67–9.
Dryer [1906]: *History of the Planetary Systems from Thales to Kepler*, 1906.
Duhem [1906]: *La Théorie Physique, Son Objet et Sa Structure*, 1905. English translation of the second (1914) edition: *The Aim and Structure of Physical Theory*, 1954.
Eccles [1964]: 'The Neurophysiological Basis of Experience', in Bunge (*ed.*): *The Critical Approach to Science and Philosophy*, 1964.
Ehrenfest [1911]: 'Welche Züge der Lichtquantenhypothese spielen in der Theorie der Wärmestrahlung eine wesentliche Rolle?', *Annalen der Physik*, **36**, pp. 91–118.
Ehrenfest [1913]: 'Zur Krise der Lichtäther-Hypothese', 1913.
Einstein [1909]: 'Über die Entwicklung unserer Anschauungen über das Wesen und die Konstitution der Strahlung', *Physikalische Zeitschrift*, **10**, pp. 817–26.
Einstein [1927]: 'Neue Experimente über den Einfluss der Erdbewegung auf die Lichtgeschwindigkeit relativ zur Erde', *Forschungen und Fortschritte*, **3**, p. 36.
Einstein [1928]: Letter to Schrödinger, 31.5.1928; published in K. Przibram (*ed.*): *Briefe Zur Wellenmechanik*, 1963.
Einstein [1931]: 'Gedenkworte auf Albert A. Michelson', *Zeitschrift für angewandte Chemie*, **44**, p. 658.

Einstein [1949]: 'Autobiographical Notes', in Schilpp (*ed.*): *Albert Einstein, Philosopher-Scientist*, **1**, pp. 2–95.
Ellis and Mott [1933]: 'Energy Relations in the β-Ray Type of Radioactive Disintegration', *Proceedings of the Royal Society of London*, Series A, **96**, pp. 502–11.
Ellis and Wooster [1927]: 'The average Energy of Disintegration of Radium E', *Proceedings of the Royal Society*, Series A, **117**, pp. 109–23.
Evans [1913]: 'The Spectra of Helium and Hydrogen', *Nature*, **92**, p. 5.
Fermi [1933]: 'Tentativo di una teoria dell emissione dei raggi "beta"', *Ricerci Scientifica*, **4(2)**, pp. 491–5.
Fermi [1934]: 'Versuch einer Theorie der β-Strahlen. I', *Zeitschrift für Physik*, **88**, pp. 161–77.
Feyerabend [1959]: 'Comments on Grünbaum's "Law and Convention in Physical Theory" ', in Feigl and Maxwell (*eds*): *Current Issues in the Philosophy of Science*, 1961, pp. 155–61.
Feyerabend [1965]: 'Reply to Criticism', in Cohen and Wartofsky (*eds*): *Boston Studies in the Philosophy of Science*, II, pp. 223–61.
Feyerabend [1968–9]: 'On a Recent Critique of Complementarity', *Philosophy of Science*, **35**, pp. 309–31 and **36**, pp. 82–105.
Feyerabend [1969]: 'Problems of Empiricism II', in Colodny (*ed.*): *The Nature and Function of Scientific Theory*, 1969.
Feyerabend [1970]: 'Against Method', *Minnesota Studies for the Philosophy of Science*, **4**, 1970.
Fowler [1912]: 'Observations of the Principal and Other Series of lines in the Spectrum of Hydrogen', *Monthly Notices of the Royal Astronomical Society*, **73**, pp. 62–71.
Fowler [1913*a*]: 'The Spectra of Helium and Hydrogen', *Nature*, **92**, p. 95.
Fowler [1913*b*]: 'The Spectra of Helium and Hydrogen', *Nature*, **92**, p. 232.
Fowler [1914]: 'Series Lines in Spark Spectra', *Proceedings of the Royal Society of London (A)*, **90**, pp. 426–30.
Fresnel [1818]: 'Lettre à François Arago sur l'Influence du Mouvement Terrestre dans quelques Phénomènes Optiques', *Annales de Chimie et de Physique*, **9**, pp. 57 ff.
Galileo [1632]: *Dialogo dei Massini Sistemi*, 1632.
Gamow [1966]: *Thirty Years that Shook Physics*, 1966.
Grünbaum [1959*a*]: 'The Falsifiability of the Lorentz-Fitzgerald Contraction Hypothesis', *British Journal for the Philosophy of Science*, **10**, pp. 48–50.
Grünbaum [1959*b*]: 'Law and Convention in Physical Theory', in Feigl and Maxwell (*eds*): *Current Issues in the Philosophy of Science*, 1961, pp. 40–155.
Grünbaum [1960]: 'The Duhemian Argument', *Philosophy of Science*, **11**, pp. 75–87.
Grünbaum [1969]: 'Can We Ascertain the Falsity of a Scientific Hypothesis?', *Studium Generale*, **22**, pp. 1061–93.
Heisenberg [1955]: 'The Development of the Interpretation of Quantum Theory', in Pauli (*ed.*): *Niels Bohr and the Development of Physics*, 1955.
Hempel [1937]: Review of Popper's [1934], *Deutsche Literaturzeitung*, 1937, pp. 309–14.
Hempel [1952]: 'Some Theses on Empirical Certainty', *The Review of Metaphysics*, **5**, pp. 620–1.
Henderson [1934]: 'The Upper Limits of the Continuous β-ray Spectra of Thorium C and C^{11}', *Proceedings of the Royal Society of London*, Series A, **147**, pp. 572–82.
Hevesy [1913]: 'Letter to Rutherford, 14.10.1913', quoted in Bohr [1963], p. XLII.
Hund [1961]: 'Göttingen, Copenhagen, Leipzig im Rückblick', in Bopp (*ed.*): *Werner Heisenberg und die Physik unserer Zeit*, Braunschweig 1961.
Jaffe [1960]: *Michelson and the Speed of Light*, 1960.
Jammer [1966]: *The Conceptual Development of Quantum Mechanics*, 1966.
Joffé [1911]: 'Zur Theorie der Strahlungserscheinungen', *Annalen der Physik*, **36**, pp. 534–52.
Juhos [1966]: 'Über die empirische Induktion', *Studium Generale*, **19**, pp. 259–72.
Keynes [1921]: *A Treatise on Probability*, 1921.

Koestler [1959]: *The Sleepwalkers*, 1959.
Konopinski and Uhlenbeck [1935]: 'On the Fermi theory of β-radioactivity', *Physical Review*, **48**, pp. 7–12.
Kramers [1923]: 'Das Korrespondenzprinzip und der Schalenbau des Atoms', *Die Naturwissenschaften*, **11**, pp. 550–9.
Kudar [1929–30]: 'Der wellenmechanische Charakter des β-Zerfalls, I–II–III', *Zeitschrift für Physik*, **57**, pp. 257–60, **60**, pp. 168–75 and 176–83.
Kuhn [1962]: *The Structure of Scientific Revolutions*, 1962.
Kuhn [1965]: 'Logic of Discovery or Psychology of Research', *this volume*, pp. 1–23.
Lakatos [1962]: 'Infinite Regress and the Foundations of Mathematics', *Aristotelian Society Supplementary Volume*, **36**, pp. 155–84.
Lakatos [1963–4]: 'Proofs and Refutations', *The British Journal for the Philosophy of Science*, **14**, pp. 1–25, 120–39, 221–43, 296–342.
Lakatos [1968*a*]: 'Changes in the Problem of Inductive Logic', in Lakatos (*ed.*): *The Problem of Inductive Logic*, 1968, pp. 315–417.
Lakatos [1968*b*]: 'Criticism and the Methodology of Scientific Research Programmes', in *Proceedings of the Aristotelian Society*, 69, pp. 149–86.
Lakatos [1970]: *The Changing Logic of Scientific Discovery*, 1970.
Lakatos [1971]: *Proofs and Refutations and Other Essays in the Philosophy of Mathematics*, 1971.
Laplace [1796]: *Exposition du Système du Monde*, 1796.
Larmor [1904]: 'On the Ascertained Absence of Effects of Motion through the Aether, in Relation to the Constitution of Matter, and on the Fitzgerald-Lorentz Hypothesis', *Philosophical Magazine*, Series 6, **7**, pp. 621–5.
Laudan [1965]: 'Grünbaum on "The Duhemian Argument" ', *Philosophy of Science*, **32**, pp. 295–9.
Leibnitz [1678]: Letter to Conring, 19.3.1678.
Le Roy [1899]: 'Science et Philosophie', *Revue de Métaphysique et de Morale*, **7**, pp. 375–425, 503–62, 706–31.
Le Roy [1901]: 'Un Positivisme Nouveau', *Revue de Métaphysique et de Morale*, **9**, pp. 138–53.
Lorentz [1886]: De l'Influence du Mouvement de la Terre sur les Phénomènes Lumineux', *Versl. Kon. Akad. Wetensch. Amsterdam*, **2**, pp. 297–358. Reprinted in Lorentz: *Collected Papers*, **4**, 1937, pp. 153–218.
Lorentz [1892*a*]: 'The Relative Motion of the Earth and the Ether', *Versl. Kon. Akad. Wetensch. Amsterdam*, **1**, pp. 74–7. Reprinted in Lorentz: *Collected Papers*, **4**, 1937, pp. 219–23.
Lorentz [1892*b*]: 'Stokes' Theory of Aberration', *Versl. Kon. Akad. Wetensch. Amsterdam*, **1**, pp. 97–103. Reprinted in Lorentz: *Collected Papers*, **4**, 1937, pp. 224–31.
Lorentz [1895]: *Versuch einer Theorie der electrischen und optischen Erscheinungen in bewegten Körpern*, 1895, § 89–92.
Lorentz [1897]: 'Concerning the Problem of the Dragging Along of the Ether by the Earth', *Versl. Kon. Akad. Wetensch. Amsterdam*, **6**, pp. 266–72. Reprinted in Lorentz: *Collected Papers*, **4**, 1937, pp. 237–44.
Lorentz [1923]: 'The Rotation of the Earth and its Influence on Optical Phenomena', *Nature*, **112**, pp. 103–4.
Lykken [1968]: 'Statistical Significance in Psychological Research', *Psychological Bulletin*, **70**, pp. 151–9.
McCulloch [1825]: *The Principles of Political Economy: With a Sketch of the Rise and Progress of the Science*, 1825.
MacLaurin [1748]: *Account of Sir Isaac Newton's Philosophical Discoveries*, 1748.
Margenau [1950]: *The Nature of Physical Reality*, 1950.
Marignac [1860]: 'Commentary on Stas' Researches on the Mutual Relations of Atomic Weights', reprinted in *Prout's Hypothesis*, Alembic Club Reprints, **20**, pp. 48–58.
Maxwell [1871]: *Theory of Heat*, 1871.

Medawar [1967]: *The Art of the Soluble*, 1967.
Medawar [1969]: *Induction and Intuition in Scientific Thought*, 1969.
Meehl [1967]: 'Theory Testing in Psychology and Physics: a Methodological Paradox', *Philosophy of Science*, **34**, pp. 103–15.
Meitner [1933]: 'Kernstruktur', in Geiger-Scheel (*eds.*): *Handbuch der Physik*, Zweite Auflage, **22/1**, pp. 118–52.
Meitner and Orthmann [1930]: 'Uber eine absolute Bestimmung der Energie der primären β-Strahlen von Radium E', *Zeitschrift für Physik*, **60**, pp. 143–55.
Michelson [1881]: 'The Relative Motion of the Earth and the Luminiferous Ether', *American Journal of Science*, Ser. 3, **22**, pp. 120–9.
Michelson [1891–2]: 'On the Application of Interference Methods to Spectroscopic Measurements, I–II', *Philosophical Magazine*, Ser. 3, **31**, pp. 338–46, and **34**, pp. 280–99.
Michelson [1897]: 'On the Relative Motion of the Earth and the Ether', *American Journal of Science*, Ser. 4, **3**, pp. 475–8.
Michelson and Gale [1925]: 'The Effect of the Earth's Rotation on the Velocity of Light', *Astrophysical Journal*, **61**, pp. 137–45.
Michelson and Morley [1887]: 'On the Relative Motion of the Earth and the Luminiferous Ether', *American Journal of Science*, Ser. 3, **34**, pp. 333–45.
Milhaud [1896]: 'La Science Rationnelle', *Revue de Métaphysique et de Morale*, **4**, pp. 280–302.
Mill [1843]: *A System of Logic, Ratiocinative and Inductive, Being a Connected View of the Principles of Evidence, and the Methods of Scientific Investigation*, 1843.
Miller [1925]: 'Ether-Drift Experiments at Mount Wilson', *Science*, **41**, pp. 617–21.
Morley and Miller [1904]: Letter to Kelvin, published in *Philosophical Magazine*, Ser. 6, **8**, pp. 753–4.
Moseley [1914]: 'Letter to Nature', *Nature*, **92**, p. 554.
Mott [1933]: "Wellenmechanik und Kernphysik', in Geiger and Scheel (*eds.*): *Handbuch der Physik*, Zweite Auflage, **24/1**, pp. 785–841.
Musgrave [1968]: 'On a Demarcation Dispute', in Lakatos and Musgrave (*eds.*): *Problems in the Philosophy of Science*, 1968, pp. 78–88.
Musgrave [1969*a*]: *Impersonal Knowledge*, Ph.D. Thesis, University of London, 1969.
Musgrave [1969*b*]: Review of Ziman's 'Public Knowledge: An Essay Concerning the Social Dimensions of Science', in *The British Journal for the Philosophy of Science*, **20**, pp. 92–4.
Nagel [1961]: *The Structure of Science*, 1961.
Nagel [1967]: 'What is True and False in Science: Medawar and the Anatomy of Research', *Encounter*, **29**, No. 3, pp. 68–70.
Nature [1913–14]: 'Physics at the British Association', *Nature*, **92**, pp. 305–9.
Neurath [1935]: 'Pseudorationalismus der Falsifikation', *Erkenntnis*, **5**, pp. 353–65.
Nicholson [1913]: 'A Possible Extension of the Spectrum of Hydrogen', *Monthly Notices of the Royal Astronomical Society*, **73**, pp. 382–5.
Pauli [1958]: 'Zur älteren und neueren Geschichte des Neutrinos', published in Pauli, *Aufsätze und Vorträge über Physik und Erkenntnistheorie*, 1961, pp. 156–80.
Pearce Williams [1968]: *Relativity Theory: Its Origins and Impact on Modern Thought*, 1968.
Peierls [1936]: 'Interpretation of Shankland's Experiment', *Nature*, **137**, p. 904.
Planck (1900*a*]: 'Über eine Verbesserung der Wienschen Spektralgleichung', *Verhandlungen der Deutschen Physikalischen Gesellschaft*, **2**, pp. 202–4; English translation in Ter Haar [1967].
Planck [1900*b*]: 'Zur Theorie des Gesetzes der Energieverteilung im Normalspektrum', *Verhandlungen der Deutschen Physikalischen Gesellschaft*, **2**, pp. 237–45; English translation in Ter Haar [1967].
Planck [1929]: 'Zwanzig Jahre Arbeit am Physikalischen Weltbild', *Physica*, **9**, pp. 193–222.

Planck [1947]: *Scientific Autobiography*, published posthumously in German in 1948, in English translation in 1950.
Poincaré [1891]: 'Les géométries non euclidiennes', *Revue des Sciences Pures et Appliquées*, **2**, pp. 769–74.
Poincaré [1902]: *La Science et l'Hypothèse*, 1902.
Polanyi [1958]: *Personal Knowledge, Towards a Post-critical Philosophy*, 1958.
Popkin [1968]: 'Scepticism, Theology and the Scientific Revolution in the Seventeenth Century', in Lakatos and Musgrave (*eds.*): *Problems in the Philosophy of Science*, 1968, pp. 1–28.
Popper [1933]: 'Ein Kriterium des empirischen Charakters theoretischer Systeme', *Erkenntnis*, **3**, pp. 426–7.
Popper [1934]: *Logik der Forschung*, 1935 (expanded English edition: Popper [1959*a*]).
Popper [1935]: 'Induktionslogik und Hypothesenwahrscheinlichkeit', *Erkenntnis*, **5**, pp. 170–2.
Popper [1940]: 'What is Dialectic?', *Mind*, N.S. **49**, pp. 403–26; reprinted in Popper [1963], pp. 312–35.
Popper [1945]: *The Open Society and its Enemies*, I–II, 1945.
Popper [1957]: 'The Aim of Science', *Ratio*, **1**, pp. 24–35.
Popper [1958]: 'Philosophy and Physics'; published in *Atti del XII Congresso Internazionale di Filosofia*, Vol. 2, 1960, pp. 363–74.
Popper [1959*a*]: *The Logic of Scientific Discovery*, 1959.
Popper [1959*b*]: 'Testability and "ad-Hocness" of the Contraction Hypothesis', *British Journal for the Philosophy of Science*, **10**, p. 50.
Popper [1963]: *Conjectures and Refutations*, 1963.
Popper [1965]: 'Normal Science and its Dangers', *this volume*, pp. 51–8.
Popper [1968*a*]: 'Epistemology without a Knowing Subject', in Rootselaar-Staal (*eds.*): *Proceedings of the Third International Congress for Logic, Methodology and Philosophy of Science*, Amsterdam, 1968, pp. 333–73.
Popper (1968*b*]: 'On the Theory of the Objective Mind', in *Proceedings of the XIV International Congress of Philosophy*, **1**, 1968, pp. 25–53.
Popper [1968*c*]: 'Remarks on the Problems of Demarcation and Rationality', in Lakatos and Musgrave (*eds.*): *Problems in the Philosophy of Science*, 1968, pp. 88–102.
Popper [1969]: 'A Realist View of Logic, Physics and History', in Yourgrau (*ed.*): *Logic, Physics and History*, 1969.
Power [1964]: *Introductory Quantum Electrodynamics*, 1964.
Prokhovnik [1967]: *The Logic of Special Relativity*, 1967.
Prout [1815]: 'On the Relation between the Specific Gravities of Bodies in their Gaseous State and the Weights of their Atoms', *Annals of Philosophy*, **6**, pp. 321–30; reprinted in *Prout's Hypothesis*, Alembic Club Reprints, **20**, 1932.
Quine [1953]: *From a Logical Point of View*, 1953.
Rabi [1961]: 'Atomic Structure', in G. M. Murphy and M. H. Shamos (*eds.*): *Recent Advances in Science*, 1956.
Reichenbach [1951]: *The Rise of Scientific Philosophy*, 1951.
Runge [1925]: 'Äther und Relativitätstheorie', *Die Naturwissenschaften*, **13**, p. 440.
Russell [1914]: *The Philosophy of Bergson*, 1914.
Russell [1943]: 'Reply to Critics', in Schilpp (*ed.*): *The Philosophy of Bertrand Russell*, 1943, pp. 681–741.
Russell [1946]: *History of Western Philosophy*, 1946.
Rutherford, Chadwick and Ellis [1930]: *Radiations from Radioactive Substances*, 1930.
Schlick [1934]: 'Über das Fundament der Erkenntnis', *Erkenntnis*, **4**, pp. 79–99; published in English in Ayer (*ed.*): *Logical Positivism*, 1959, pp. 209–27.
Schrödinger [1958]: 'Might perhaps Energy be merely a Statistical Concept?', *Il Nouvo Cimento*, **9**, pp. 162–70.

Shankland [1936]: 'An Apparent Failure of the Photon Theory of Scattering', *Physical Review*, **49**, pp. 8–13.
Shankland [1964]: 'Michelson-Morley Experiment', *American Journal of Physics*, **32**, pp. 16–35.
Soddy [1932]: *The Interpretation of the Atom*, 1932.
Sommerfeld [1916]: 'Zur Quantentheorie der Spektrallinien', *Annalen der Physik*, **51**, pp. 1–94 and 125–67.
Stebbing [1914]: *Pragmatism and French Voluntarism*, 1914.
Stegmüller [1966]: 'Explanation, Prediction, Scientific Systematization and Non-Explanatory Information', *Ratio*, **8**, pp. 1–24.
Stokes [1845]: 'On the Aberration of Light', *Philosophical Magazine*, Third Series, **27**, pp. 9–15.
Stokes [1846]: 'On Fresnel's Theory of the Aberration of Light', *Philosophical Magazine*, Third Series, **28**, pp. 76–81.
Synge [1952–4]: 'Effects of Acceleration in the Michelson-Morley Experiment', *The Scientific Proceedings of the Royal Dublin Society*, New Series, **26**, pp. 45–54.
Ter Haar [1967]: *The Old Quantum Theory*, 1967.
Thomson [1929]: 'On the Waves associated with β-rays, and the Relation between Free Electrons and their Waves', *Philosophical Magazine*, Seventh Series, **7**, pp. 405–17.
Toulmin [1967]: 'The Evolutionary Development of Natural Science', *American Scientist*, **55**, pp. 456–71.
Treiman [1959]: 'The Weak Interactions', *Scientific American*, **200**, pp. 72–84.
Truesdell [1960]: 'The Program toward Rediscovering the Rational Mechanics in the Age of Reason', *Archive of the History of Exact Sciences*, **1**, pp. 3–36.
Uhlenbeck and Goudsmit [1925]: 'Ersetzung der Hypothese vom unmechanischen Zwang durch eine Forderung bezüglich des inneren Verhaltens jedes einzelnen Electrons', *Die Naturwissenschaften*, **13**, pp. 953–4.
van der Waerden [1967]: *Sources of Quantum Mechanics*, 1967.
Watkins [1957]: 'Between Analytic and Empirical', *Philosophy*, **32**, pp. 112–31.
Watkins [1958]: 'Influential and Confirmable Metaphysics', *Mind*, N.S. **67**, pp. 344–65.
Watkins [1960]: 'When are Statements Empirical?', *British Journal for the Philosophy of Science*, **10**, pp. 287–308.
Watkins [1968]: 'Hume, Carnap and Popper', in Lakatos (*ed.*): *The Problem of Inductive Logic*, 1968, pp. 271–82.
Whewell [1837]: *History of the Inductive Sciences, from the Earliest to the Present Time*. Three volumes, 1837.
Whewell [1840]: *Philosophy of the Inductive Sciences, Founded upon their History*. Two volumes, 1840.
Whewell [1851]: 'On the Transformation of Hypotheses in the History of Science', *Cambridge Philosophical Transactions*, **9**, pp. 139–47.
Whewell [1858]: *Novum Organon Renovatum*. Being the second part of the philosophy of the inductive sciences. Third edition, 1858.
Whewell [1860]: *On the Philosophy of Discovery, Chapters Historical and Critical*, 1860.
Whittaker [1947]: *From Euclid to Eddington*, 1947.
Whittaker [1953]: *History of the Theories of Aether and Electricity*, Vol. II, 1953.
Wisdom [1963]: 'The Refutability of "Irrefutable" Laws', *The British Journal for the Philosophy of Science*, **13**, pp. 303–6.
Wu [1966]: 'Beta Decay', in *Rendiconti della Scuola Internazionale di Fisica*, "Enrico Fermi", *XXXII* Corso.
Wu and Moskowski [1966]: *Beta Decay*, 1966.